Karl R. Popper

CONOCIMIENTO OBJETIVO

Un enfoque evolucionista

Tecnos

Karl R. Popper

CONOCIMIENTO OBJETIVO

Un enfoque evolucionista

CUARTA EDICION

tecnos

Los derechos para la versión castellana de la obra
Objective Knowledge (1.ª ed., 1972)
© Karl R. Popper, 1972
son propiedad de EDITORIAL TECNOS (GRUPO ANAYA, S. A.),
por acuerdo con THE CLARENDON PRESS OXFORD.

Traducción por:
CARLOS SOLÍS SANTOS

Diseño de cubierta:
Joaquín Gallego

1.ª edición, 1974
2.ª edición, 1982
3.ª edición, 1988
4.ª edición, 1992
Reimpresión, 2001

Reservados todos los derechos. El contenido de esta obra está protegido por la Ley, que establece penas de prisión y/o multas, además de las correspondientes indemnizaciones por daños y perjuicios, para quienes reprodujeren, plagiaren, distribuyeren o comunicaren públicamente, en todo o en parte, una obra literaria, artística o científica, o su transformación, interpretación o ejecución artística fijada en cualquier tipo de soporte o comunicada a través de cualquier medio, sin la preceptiva autorización.

© EDITORIAL TECNOS (GRUPO ANAYA, S. A.), 2001
Juan Ignacio Luca de Tena, 15 - 28027 Madrid
ISBN: 84-309-0488-3
Depósito Legal: M-8.776-2001

Printed in Spain. Impreso en España por EFCA, S.A.

Dedicado a Alfred Tarski

PREFACIO

El conocimiento humano es sin duda el mayor milagro de nuestro universo. Constituye un problema cuya solución no es inminente y no creo en absoluto que este volumen sea siquiera un pequeño paso hacia dicha solución. No obstante, espero haber conseguido iniciar una discusión que ha estado empantanada en cuestiones preliminares durante tres siglos.

Desde Descartes, Hobbes, Locke y su escuela, en la que hay que contar no sólo a David Hume, sino también a Thomas Reid, la teoría del conocimiento humano ha sido en gran medida subjetivista, el conocimiento se consideraba como un tipo especialmente seguro de creencia humana y el conocimiento científico, como un tipo especialmente seguro de conocimiento humano.

Los ensayos de este libro rompen con una tradición que puede ser retrotraída hasta Aristóteles: la tradición de esa teoría del conocimiento de sentido común. Soy un gran admirador del sentido común que considero sin duda como esencialmente autocrítico. Pero mientras que estoy dispuesto a defender hasta el final la verdad esencial del realismo de sentido común, *considero que la* teoría del conocimiento de sentido común *es un desatino subjetivista. Este desatino ha dominado la Filosofía occidental. He intentado erradicarlo y reemplazarlo por una teoría objetiva del conocimiento esencialmente conjetural. Tal vez esta pretensión sea osada, pero no voy a disculparme por ello.*

Sin embargo, considero un deber disculparme por ciertas repeticiones: he dejado los diversos capítulos, publicados o no anteriormente, casi en el mismo estado en que fueron escritos, aun cuando se superpusiesen parcialmente. Precisamente por esta razón hablo en los capítulos 3 y 4 de "primer", "segundo" y "tercer mundo", aunque ahora, siguiendo una sugerencia de Sir John Eccles en su libro Facing Reality, *prefiero hablar de "mundo 1", "mundo 2" y "mundo 3", como en el capítulo 2.*

KARL R. POPPER

Penn, Buckinghamshire
24 de julio de 1971

AGRADECIMIENTOS

Estoy en deuda con David Miller, Arne F. Petersen, Jeremy Shearmur y, sobre todo, con mi esposa, por su ayuda paciente e infatigable.

K. R. P.

INDICE

PREFACIO ... ... *Pág.* 9

1. EL CONOCIMIENTO COMO CONJETURA: MI SOLUCIÓN AL PROBLEMA DE LA INDUCCIÓN 15

1. El problema de la inducción del sentido común, 16.—2. Los dos problemas de la inducción de Hume, 17.—3. Consecuencias importantes de los resultados de Hume, 18.—4. Mi manera de enfocar el problema de la inducción, 19.—5. El problema lógico de la inducción: replanteamiento y solución, 20.—6. Comentarios a mi solución del problema lógico, 21.—7. La preferencia entre teorías y búsqueda de la verdad, 25.—8. Corroboración: los méritos de la improbabilidad, 29.—9. Preferencia pragmática, 32.—10. Bases de mi replanteamiento del problema psicológico de la inducción de Hume, 34.—11. Replanteamiento del problema psicológico de la inducción, 36.—12. El problema tradicional de la inducción y la invalidez de todos los principios o reglas de inducción, 37.—13. Más allá de los problemas de inducción y demarcación, 39.

2. LAS DOS CARAS DEL SENTIDO COMÚN: ARGUMENTOS EN PRO DEL REALISMO DEL SENTIDO COMÚN Y EN CONTRA DE LA TEORÍA DEL CONOCIMIENTO DEL SENTIDO COMÚN 41

1. Defensa de la filosofía, 41.—2. Punto de partida incierto: sentido común y crítica, 42.—3. Diferencias con otros enfoques, 43.—4. Realismo, 45.—5. Argumentos en favor del realismo, 45.—6. Consideraciones sobre la verdad, 51.—7. Contenido, contenido de verdad y contenido de falsedad, 53.—8. Consideraciones sobre la verosimilitud, 58.—9. Verosimilitud y la búsqueda de la verdad, 59.—10. Verdad y verosimilitud como objetivos, 62.—11. Comentarios en torno a las nociones de verdad y verosimilitud, 63.—12. La equivocada teoría del conocimiento del sentido común, 65.—13. Crítica de la teoría del conocimiento del sentido común, 67.—14. Crítica de la teoría subjetivista del conocimiento, 69.—15. Carácter pre-darwinista de la teoría del conocimiento del sentido común, 70.—16. Bosquejo de una epistemología evolucionista, 71. 17. Conocimiento básico y problemas, 74.—18. Todo conocimiento —incluso las observaciones— está impregnado de teoría, 74.—19. Examen retrospectivo de la epistemología subjetivista, 75.—20. Conocimiento en sentido objetivo, 76.—21. La búsqueda de la certeza y la debilidad fundamental de la teoría del conocimiento del sentido común, 77.—22. Consideraciones analíticas sobre la certeza, 80.—23. El método de la ciencia, 83.—24. Discusión crítica, preferencia racional y el problema de la analiticidad de nuestras elecciones y predicciones, 84.—25. Ciencia: el aumento del conocimiento mediante crítica e ingenio, 86.—*Reflexión retrospectiva sobre la inducción*, 86.—26. Los problemas de la causalidad y la inducción en Hume, 86.—27. Por qué el problema lógico de la inducción de Hume es más profundo que su problema de la causalidad, 91.—

28. Intervención kantiana: conocimiento objetivo, 92.—29. La solución de la paradoja de Hume: restauración de la racionalidad, 94.—30. Embrollos relacionados con el problema de la inducción, 96.—31. ¿Qué es lo que queda del problema erróneo de la justificación de la inducción?, 97.—32. Escepticismo dinámico: confrontación con Hume, 99.—33. Análisis de un argumento sacado de la improbabilidad de los accidentes, 101.—34. Sumario: Filosofía crítica del sentido común, 103.

3. EPISTEMOLOGÍA SIN SUJETO COGNOSCENTE 106

1. Tres tesis sobre epistemología y el tercer mundo, 106.—2. Aproximación biológica al tercer mundo, 111.—3. La objetividad y autonomía del tercer mundo, 114.—4. Lenguaje, crítica y el tercer mundo, 117.—5. Consideraciones históricas, 120.—6. Evaluación y crítica de la epistemología de Brouwer, 126.—7. Subjetivismo en lógica, teoría de la probabilidad y física, 135.—8. Lógica y biología de la investigación, 138.—9. Descubrimiento, humanismo y autotrascendencia, 141.—Bibliografía selecta, 144.

4. SOBRE LA TEORÍA DE LA MENTE OBJETIVA 147

1. Pluralismo y la tesis de los tres mundos, 147.—2. Las relaciones causales entre los tres mundos, 148.—3. La objetividad del tercer mundo, 150.—4. El tercer mundo como producto del hombre, 151.—5. El problema de la comprensión, 154.—6. Procesos psicológicos de pensamiento y objetos del tercer mundo, 156.—7. Comprensión y resolución de problemas, 158.—8. Un ejemplo muy trivial, 160.—9. Un caso de comprensión histórica objetiva, 162.—10. El valor de los problemas, 170.—11. La comprensión ("hermenéutica") en las humanidades, 173.—12. Comparación con el método de la evocación subjetiva de Collingwood, 176.

5. EL OBJETO DE LA CIENCIA 180
Bibliografía selecta, 191.—Nota bibliográfica, 192.

6. SOBRE NUBES Y RELOJES 193
Aproximación al problema de la racionalidad y libertad del hombre, 193.

7. LA EVOLUCIÓN Y EL ÁRBOL DEL CONOCIMIENTO 236

1. Algunas consideraciones en torno a los problemas y desarrollo del conocimiento, 237.—2. Consideraciones en torno a los métodos de la biología y en especial de la teoría de la evolución, 244.—3. Una conjetura: "dualismo genético", 250.—*Suplemento.* El prometedor monstruo comportamental, 256.

8. CONSIDERACIÓN REALISTA DE LA LÓGICA, LA FÍSICA Y LA HISTORIA 260

1. Realismo y pluralismo: reducción versus emergencia, 263.—2. Pluralismo y emergencia en la historia, 269.—3. Realismo y subjetivismo en física, 273.—4. Realismo en lógica, 276.

9. COMENTARIOS FILOSÓFICOS EN TORNO A LA TEORÍA DE LA VERDAD DE TARSKI. 288
Suplemento. Nota sobre la definición de verdad de Tarski, 302.

APÉNDICE. El cubo y el reflector: dos teorías acerca del conocimiento 307

INDICE DE AUTORES 327

INDICE DE MATERIAS 331

1. EL CONOCIMIENTO COMO CONJETURA: MI SOLUCION AL PROBLEMA DE LA INDUCCION *

> El aumento de la irracionalidad durante el siglo diecinueve y lo ocurrido en el veinte es el resultado natural de la destrucción del empirismo por parte de Hume.—BERTRAND RUSSELL.

Creo que he resuelto uno de los mayores problemas filosóficos: el problema de la inducción. (He debido hallar la solución hacia el año 1927, más o menos [1].) Esta solución, que ha resultado ser de una gran fecundidad, me ha permitido resolver muchos otros problemas filosóficos.

Sin embargo, pocos filósofos aceptarán la tesis de que he resuelto el problema de la inducción. Pocos filósofos se han tomado la molestia de estudiar —e incluso de criticar— mis opiniones sobre este problema, o se han enterado de mis investigaciones. Muy recientemente, se han publicado varios libros sobre el tema que no citan ninguno de mis trabajos, aunque la mayoría acusan la influencia, muy lejana, de mis ideas. Incluso obras que las recogen explícitamente, me atribuyen opiniones que jamás he sostenido y me critican basándose en una total incompresión de lo que he escrito, o en argumentos inadecuados. En este capítulo intento explicar de nuevo mis puntos de vista dando cabal respuesta a mis críticos.

Mis dos primeras publicaciones sobre la inducción fueron una nota en *Erkenntnis* de 1933 [2], en la que presentaba brevemente la formulación y solución del problema y mi libro *Logik der Forschung (L. d. F.)* de 1934 [3]. En ambas obras, mis ideas aparecían condensadas. Esperaba con cierto

---

* Este capítulo se publicó por primera vez en la *Revue internationale de Philosophie*, 25e année, núm. 95-6, fasc. 1-2.

[1] Ya en el invierno de 1919-20 había yo formulado y resuelto el problema de la demarcación entre la ciencia y lo que no es ciencia, si bien entonces no creía que mereciese la pena publicarlo. Pero, después de resolver el problema de la inducción, descubrí una conexión interesante entre ambos problemas, lo cual me hizo pensar que el de la demarcación era importante. Comencé a trabajar en el problema de la inducción en 1923 y di con la solución hacia el año 1927. Véanse a este respecto las consideraciones autobiográficas que aparecen en mi libro, *Conjectures and Refutations*, (*C. & R.*, para abreviar) capítulos I y II. [Traducción castellana de Néstor Míguez, *El Desarrollo del Conocimiento Científico. Conjeturas y Refutaciones*, Buenos Aires, Paidos, 1967.]

[2] "Ein Kriterium des empirischen Charakters theoretischer Systeme", *Erkenntnis*, 3, 1933, pág. 426 y siguientes.

[3] *Logik der Forschung*, Julius Springer Verlag, Viena, 1934 (para abreviar, *L. d. F.*). Cf. *The Logic of Scientific Discovery*, Hutchison, Londres, 1959 (para abreviar, *L. Sc. D.*). [Traducción castellana de Víctor Sánchez de Zavala, *La Lógica de la Investigación Científica*, Madrid, Tecnos, 1962.]

optimismo, que mis lectores descubriesen, con ayuda de algunas sugerencias históricas, por qué era decisiva mi peculiar *reformulación* del problema, ya que creo que lo que hizo posible la solución del problema tradicional fue dicha reformulación.

Al hablar del *problema filosófico tradicional de la inducción* (que llamaré "*Tr*") me refiero a formulaciones como las siguientes:

Tr ¿Cómo se justifica la creencia de que el futuro será (en gran medida) como el pasado? O, tal vez, ¿cómo se justifican las inferencias inductivas?

Estas preguntas están mal formuladas por diversas razones. La primera, por ejemplo, *supone* que el futuro será como el pasado. Personalmente, considero un error suponer tal cosa, a menos que la palabra "como" se tome en un sentido tan flexible que haga que dicha suposición sea vacía e inocua. La segunda formulación supone la existencia de inferencias inductivas y de *reglas* para obtenerlas, lo que significa, una vez más, hacer una suposición acrítica que me parece equivocada. Pienso, por tanto, que ambas formulaciones son sencillamente acríticas y lo mismo podría decirse de muchas otras. Por tanto, me he impuesto fundamentalmente la tarea de formular una vez más el *problema que está detrás, a mi parecer,* de lo que he llamado el problema filosófico tradicional de la inducción.

Las formulaciones, ya tradicionales, datan de fecha histórica reciente: surgen de la crítica de Hume a la inducción y de su impacto sobre la teoría del conocimiento del sentido común.

Volveré a discutir con mayor detalle estas formulaciones tras presentar, en primer lugar, el punto de vista del sentido común, la opinión de Hume y, finalmente, mis propias reformulaciones y soluciones del problema.

1. El problema de la inducción del sentido común

La teoría del conocimiento del sentido común, que he apodado "la teoría de la mente como un cubo", queda perfectamente recogida en la frase "nada hay en el intelecto que no haya pasado antes por los sentidos". (Ya he intentado mostrar que Parménides fue el primero en formular este punto de vista. Dicho en tono satírico: la mayor parte de los mortales no tienen nada en sus falibles intelectos que no haya pasado antes por sus falibles sentidos [4].)

A pesar de todo, *esperamos cosas* y *creemos* con firmeza en *ciertas regularidades* (leyes de la naturaleza, teorías), lo que nos lleva al problema de la inducción del sentido común (que llamaré "*Cs*"):

---

[4] Véase *Conjectures and Refutations* (*C. & R.*), Apéndice 8 a la tercera edición, 1969, especialmente las págs. 408-12. [La versión castellana carece de apéndice 8, pues está hecha sobre la segunda edición de 1965.]

Cs. ¿Cómo pueden haber surgido estas espectativas y creencias?

La respuesta del sentido común consiste en decir que estas creencias surgen en virtud de *reiteradas* observaciones hechas en el pasado: creemos que el sol saldrá mañana porque así ha ocurrido en el pasado.

Desde el punto de vista del sentido común se da por supuesto, sin más problemas, que nuestras creencias en regularidades se justifican mediante esas observaciones reiteradas responsables de su génesis. (Lo que los filósofos desde Aristóteles y Cicerón han llamado "*epagoge*" o "*inducción*" [5] es esa génesis *cum* justificación, debidas ambas a la reiteración.)

2. Los dos problemas de la inducción de Hume

Hume estaba interesado por la condición del *conocimiento* humano o, como él diría, por el problema de si nuestras creencias (o, al menos, algunas de ellas) se pueden *justificar* con razones suficientes [6]. Planteó dos preguntas, una lógica (H_L) y otra psicológica (H_{Ps}), con la característica importante de que sus respuestas chocan entre sí de algún modo.

La pregunta lógica es la siguiente [7]:

H_L ¿Cómo se justifica que, partiendo de casos (reiterados) de los que tenemos experiencia, lleguemos mediante el razonamiento a otros casos (conclusiones) de los que no tenemos experiencia?

La respuesta de Hume a H_L consiste en negar que haya alguna justificación, por grande que sea el número de repeticiones. También mostró que la situación lógica sigue siendo *exactamente la misma* cuando ponemos la palabra "probable" después de "conclusiones" o cuando sustituimos las palabras "a casos" por "a la *probalidad* de casos".

---

[5] Cicerón, *Topica*, X, 42; cf. *De Inventione*, Libro I, xxxi. 51 a xxxv. 61.

[6] Véase David Hume, *Enquiry Concerning Human Understanding*, ed. L.A. Selby-Bigge, Oxford, 1927, Sección V, parte I, pág. 56. (Cf. *C. &. R.*, pág. 21 [cf. pág. 30 de la versión castellana citada].)

[7] Hume, *Treatise on Human Nature*, ed. Selby-Bigge, Oxford, 1888, 1960, Libro I, parte III, sección VI, pág. 91; Libro I, parte III, sección XII, pág. 139 [Traducción castellana de Vicente Viqueira, *Tratado de la Naturaleza Humana*, Madrid, Calpe 1923; págs. 147 y 215 del primer volumen.] Véase también, Kant, *Prolegómena*, pág. 14 y sigs. [Traducción castellana de Julián Besteiro, Madrid, Aguilar 1954, páginas 40 y siguientes], donde llama "Problema de Hume" al problema de la existencia de enunciados válidos *a priori*. Que yo sepa, fue el primero en llamar al problema de la inducción "Problema de Hume", aunque sin duda debe de haber habido otros. Así lo hice constar en "Ein Kriterium des empirischen Characters Theoretischer Systeme", *Erkenntnis, 3*, 1933, págs. 426 y sigs. y en *L. d. F.*, sección 4, pág. 7, donde escribí: "Si, siguiendo a Kant, llamamos 'problema de Hume' al de la inducción, deberíamos designar al de la demarcación como 'Problema de Kant'". [Cf. la traducción castellana citada, pág. 34.] Esta breve indicación (sostenida en unas pocas consideraciones mías, como la de la pág. 29 de *L. Sc. D.* [*Ibid*, pág. 29], en el sentido de que Kant consideraba el principio de la inducción como "válido *a priori*") sugería una interpretación histórica importante de las relaciones entre Kant, Hume y el problema de la inducción. Véase también el capítulo 2, págs. 85 y sigs., y 93 de este volumen, donde se discuten más extensamente estos problemas.

La pregunta psicológica es la siguiente [8]:

H_{PS} ¿Por qué, a pesar de todo, las personas razonables esperan y *creen* que los casos de los que no tienen experiencia van a ser semejantes a aquellos de los que tienen experiencia? Es decir, ¿por qué confiamos tanto en las experiencias que tenemos?

La respuesta de Hume se centra en la "costumbre o hábito"; es decir, porque estamos condicionados por las *repeticiones* y el mecanismo de asociación de ideas, mecanismo sin el cual, dice Hume, difícilmente sobreviviríamos.

3. Consecuencias importantes de los resultados de Hume

A causa de estos resultados, Hume, una de las mentes más racionales que haya habido nunca, se convirtió en un escéptico a la vez que en un creyente: un creyente en una epistemología irracionalista. La conclusión de que la reiteración carece de todo valor como argumento, aunque domina nuestra vida cognitiva o nuestro "entendimiento", le condujo a afirmar que la argumentación o la razón desempeñan en él un papel secundario. El entendimiento queda desenmascarado y muestra que es no ya del mismo carácter que las creencias, sino del mismo carácter que las creencias indefendibles racionalmente; es *una fe irracional* [9].

De la solución que doy al problema de la inducción no se deriva esta conclusión irracionalista, como quedará claro en la próxima sección, así como en las secciones 10 y 11.

Esta misma conclusión fue enunciada por Russell de un modo más enérgico y desesperado en el capítulo dedicado a Hume de su libro *A History of Western Philosophy*, publicado en 1946 (treinta y cuatro años después de su libro *Problems of Philosophy* en el que enunciaba con perfecta claridad el problema de la inducción sin hacer referencia a Hume) [10]. Russell dice sobre el tratamiento de la inducción que hace Hume: "La filosofía de Hume... representa la bancarrota de la racionalidad del siglo dieciocho" y, "por eso es importante averiguar si existe una respuesta a Hume en el seno de una filosofía total o fundamentalmente *empírica*. En caso contrario,

---

[8] Véase el *Treatise*, págs. 91 y 139. [págs. 147 y 215 del tomo primero de la edición castellana.]

[9] A partir de Hume, muchos inductivistas desilusionados se han hecho irracionalistas (así como también muchos marxistas desilusionados).

[10] El nombre de Hume no aparece en el capítulo VI ("Sobre la Inducción") del libro de Russell *The Problems of Philosophy* (1912 y muchas otras reimpresiones) [Hay traducción castellana de Joaquín Xirau, Barcelona, Nueva Colección Labor 1970.] La referencia más cercana aparece en el capítulo VIII ("Cómo es posible el conocimiento *apriorístico*") donde Russell afirma que Hume "dedujo la proposición mucho más dudosa, según la cual nada puede ser conocido *a priori* sobre la conexión de la causa y el efecto". Sin duda las expectativas causales tienen una base innata: son psicológicamente *a priori* en el sentido de que son previas a la experiencia. Pero esto no significa que sean válidas a *priori*. Véase *C. & R.* págs. 47-8 [traducción castellana, página 59].

no hay diferencia intelectual entre razón y locura. El lunático que cree ser un huevo escalfado ha de ser condenado únicamente porque está en minoría..."

Russell llega a decir que si se rechaza la inducción (o el principio de inducción), "todo intento de llegar a leyes científicas generales a partir de observaciones particulares es falaz, con lo que el escepticismo de Hume es insoslayable para un empirista" [11].

Por tanto, Russell subraya la contradicción existente entre la respuesta de Hume a H_L y (*a*) la racionalidad, (*b*) el empirismo y (*c*) los procedimientos científicos.

Si se acepta mi solución al problema de la inducción, todas estas dificultades desaparecen, como se verá claramente en la sección 4 y en las secciones 10 a 12. No hay contradicción entre mi teoría de la no-inducción, por un lado, y la racionalidad, el empirismo o los procedimientos científicos, por otro.

4. Mi manera de enfocar el problema de la inducción

(1) Considero de la mayor importancia la distinción, implícita en el tratamiento de Hume, entre el problema lógico y el psicológico, aunque no me satisface lo que entiende Hume por lo que llamo "lógico". Describe, con toda claridad, procesos de *inferencia válida,* pero los considera *procesos mentales* "racionales".

Por el contrario, al tratar de cualquier tipo de problemas *lógicos,* una de las maneras principales que tengo de plantear la cuestión consiste en traducir a términos objetivos todos los términos subjetivos o psicológicos, especialmente "creencia", etc. Así, en vez de hablar de "creencia", hablo, por ejemplo, de "enunciado" o de "teoría explicativa", en lugar de hablar de "impresión", hablo de "enunciado observacional" o "enunciado contrastador" [*test statement*] y en lugar de hablar de "justificación de una creencia", hablo de "justificación de la pretensión de que una teoría sea verdadera", etc.

Este modo de decir las cosas de una manera objetiva, lógica o "formal" se puede aplicar a H_L, pero no a H_{Ps}. No obstante:

(2) Una vez resuelto el problema lógico, H_L, la solución se aplica al psicológico, H_{Ps}, según el siguiente *principio de transferencia*: lo que es verdad en el dominio de la lógica, lo es también en el de la psicología. (También hay un principio análogo de gran utilidad en lo que se suele llamar "método científico", así como en la historia de la ciencia: lo que es verdad en el dominio de la lógica, lo es también en el método científico y en la historia de la ciencia.) Admito que esto constituye una

---

[11] Las citas son del libro de Bertrand Russell *A History of Western Philosophy,* Londres 1946, págs. 698 y sigs. (el subrayado es mío). [*Historia de la Filosofía Occidental,* Buenos Aires, serie mayor de Espasa Calpe.]

conjetura un tanto arriesgada en psicología del pensamiento o de los procesos cognitivos.

(3) Como se verá claramente, el principio de transferencia garantiza la eliminación del irracionalismo de Hume. Si puedo resolver el problema de la inducción, incluyendo H_{Ps}, sin violar el principio de transferencia, no habrá contradicción entre la lógica y la psicología y, por tanto, se evitará la conclusión de que nuestro conocimiento es irracional.

(4) Este programa, unido a la solución que da Hume de H_L, implica muchas más cosas sobre la relación lógica entre teorías científicas y observaciones que las que señala H_L.

(5) Uno de los principales resultados obtenidos es el siguiente: puesto que Hume está en lo cierto al señalar que desde un punto de vista *lógico* no existe inducción por repetición, en virtud del principio de transferencia, tampoco puede haber tal cosa en *psicología* (o en el método científico o en historia de la ciencia). La idea de inducción por repetición debe achacarse a un error, una especie de ilusión óptica. Resumiendo: ***no hay inducción por repetición.***

5. El problema lógico de la inducción: replanteamiento y solución

De acuerdo con lo que acabo de decir en el punto (2) de la sección precedente, he de replantear el H_L de Hume en un lenguaje objetivo y lógico.

A este fin, sustituiré la expresión de Hume "casos de los que tenemos experiencias" por "enunciados contrastadores", es decir, enunciados singulares que describen sucesos observables ("enunciados observacionales" o "enunciados básicos") y "casos de los que no tenemos experiencia" por "teorías explicativas universales".

Mi formulación del problema lógico de la inducción de Hume es la siguiente:

L_1 ¿Se puede justificar la pretensión de que una teoría explicativa universal sea verdadera mediante "razones empíricas", es decir, suponiendo la verdad de ciertos enunciados contrastadores u observacionales (los cuales, hay que decirlo, están "basados en la experiencia")?

Mi respuesta es como la de Hume: No, no podemos; ningún conjunto de enunciados contrastadores verdaderos podrá justificar la pretensión de que una teoría explicativa universal es verdadera [12].

Pero hay un segundo problema lógico, L_2, que constituye una generalización de L_1, a partir del cual se obtiene sustituyendo sencillamente las palabras "es verdadera" por "es verdadera o falsa":

L_2 ¿Se puede justificar la pretensión de que una teoría explicativa universal sea verdadera o falsa mediante "razones empíricas"? Es decir,

[12] Una teoría explicativa va necesariamente más allá de un conjunto finito de enunciados contrastadores; lo mismo se puede decir incluso de una ley con un grado reducido de universalidad.

suponiendo que los enunciados contrastadores sean verdaderos, ¿pueden ellos justificar la pretensión de que una teoría universal sea verdadera o la de que sea falsa?

A esto respondo positivamente: Sí, *suponiendo que los enunciados contrastadores sean verdaderos, basándonos en ellos podemos a veces justificar la pretensión de que una teoría explicativa universal es falsa.*

La importancia de esta respuesta se ve cuando reflexionamos sobre la situación problemática que da lugar al problema de la inducción. Pienso en la situación que se nos presenta cuando tenemos a la vista *varias teorías explicativas* que se ofrecen como otras tantas soluciones de algún problema de explicación (por ejemplo, un problema científico) y cuando debemos, o al menos deseamos, elegir entre ellos. Como hemos visto, Russell dice que sin resolver el problema de la inducción resulta imposible *decidir entre* una (buena) teoría científica y una (mala) obsesión de un demente. También Hume pensaba en teorías alternativas. "Supóngase (escribe) que una persona... establece proposiciones a las cuales no puedo asentir... que la plata es más fusible que el plomo o el mercurio más pesado que el oro..." [13]

Esta situación problemática de elegir entre varias teorías sugiere un tercer modo de formular el problema de la inducción:

L_3 Dadas varias teorías universales rivales, ¿es posible *preferir* unas a otras por lo que respecta a su verdad o falsedad, justificándolo mediante "razones empíricas"?

La respuesta a L_3 es obvia a la luz de la solución dada a L_2: Sí, a veces se puede, si hay suerte, ya que puede ocurrir que nuestros enunciados constrastadores refuten algunas —aunque no todas— de las teorías rivales y, puesto que buscamos una teoría verdadera, preferiremos aquella cuya falsedad no haya sido demostrada.

6. Comentarios a mi solución del problema lógico

(1) Según mis reformulaciones, la cuestión central del problema lógico de la inducción es la validez (verdad o falsedad) de las leyes universales *por respeto a ciertos enunciados "dados"*. Dejo de lado el problema de "cómo determinamos la verdad o la falsedad de los enunciados contrastadores", es decir, de las descripciones singulares de sucesos observables. Sugiero que este último problema no debe ser considerado como parte del problema de la inducción, puesto que la pregunta de Hume es si se puede o no justificar el paso de los "casos" experimentados a los no experimentados [14]. Que yo sepa, ni Hume ni ninguna de las demás personas que se han ocupado del tema antes que yo, ha pasado de estas preguntas a las *siguientes*: ¿Podemos dar por supuestos los "casos experimentados"? ¿Son realmente previos a las teorías? Aunque mi solución al problema de la

[13] Hume, *Treatise*, pág. 95 [pág. 159 del volumen I de la traducción castellana].
[14] *Op cit.*, pág. 91 [*trad. cit.*, pág. 150 del tomo I].

inducción me hizo ver la existencia de otros problemas, estas preguntas van más allá del problema original. (Esto es obvio si consideramos qué tipo de cosas buscaban los filósofos cuando intentaban resolver el problema de la inducción: si se pudiera encontrar un "principio de inducción" que permitiese *derivar leyes universales partiendo de enunciados singulares* y si se pudiera sostener su pretensión de verdad, entonces el problema de la inducción podría considerarse resuelto.)

(2) L_1 constituye un intento de traducir el problema de Hume a un modo objetivo de hablar. La única diferencia es que Hume habla de *casos* (singulares) futuros de los que no tenemos experiencia —es decir, de expectativas— mientras que L_1 habla de leyes universales o teorías. Propongo este cambio al menos por tres razones. En primer lugar, desde un punto de vista lógico, "casos" hace referencia a una ley universal (o al menos a una función de enunciando que puede ser universalizada). En segundo lugar, nuestro método usual de razonar de unos "casos" a otros se lleva a cabo con la ayuda de teorías universales. De este modo, pasamos del problema de Hume al *problema de la validez de las teorías universales* (su verdad o falsedad). En tercer lugar, deseo conectar, como hace Russell, el problema de la inducción con *las leyes universales o teorías científicas*.

(3) Mi respuesta negativa a L_1 ha de interpretarse en el sentido de que debemos considerar *todas las leyes o teorías como hipótesis o conjeturas;* es decir, como suposiciones.

Aunque actualmente este punto de vista ha logrado imponerse [18], esta situación no se ha logrado hasta después de muchos años. Gilbert Ryle, por ejemplo, ha atacado la idea en un artículo suyo de 1937 [19], por lo demás excelente. Arguye (p. 36) que es incorrecto afirmar "que todas las proposiciones generales de la ciencia... sean meras hipótesis". Emplea la palabra "hipótesis" en el mismo sentido en que siempre la he usado yo: es decir, en el sentido de una "proposición... cuya verdad es simplemente una conjetura" (loc. cit.). Alega en contra de mi tesis que "a menudo tenemos todas las garantías para sentirnos seguros de la ley expresada por una proposición" (p. 38). También sostiene que algunas proposiciones generales están firmemente "establecidas": "Estas se denominan 'leyes' y no 'hipótesis'."

Esta posición de Ryle era casi la norma "establecida" cuando escribía *L. d. F.* y aún sigue vigente en gran medida. Me opuse a ella en primer lugar gracias a la teoría gravitatoria de Einstein. En efecto, nunca ha habido una teoría tan firmemente "establecida" como la de Newton y es poco probable que vuelva a haber otra semejante y, sin embargo, la teoría de Einstein, piénsese lo que se quiera de ella, sin duda nos obliga considerar la teoría de Newton como una "mera" hipótesis o conjetura.

Un segundo caso es el descubrimiento que hizo Urey en 1931 del

---

[18] Véase la reflexión inicial de Stove en *Australas. Journ. of Philos.* 38, 1960, página 173.

[19] Véase *Arist Soc. Supplementary volume*, 16, 1937, págs. 36-62.

deuterio y el agua pesada. En aquella época, el agua, el hidrógeno y el oxígeno eran las sustancias químicas mejor conocidas, constituyendo los pesos atómicos del hidrógeno y el oxígeno las normas mismas de toda medida química. Había una teoría sobre cuya verdad *todo* químico se hubiera jugado el cuello, al menos hasta la conjetura de Soddy de 1910, sobre los isótopos y, de hecho, hasta mucho después. Mas he aquí de Urey dio con una refutación (que a la vez constituyó una corroboración de la teoría de Bohr).

Esta situación me hizo examinar más de cerca las restantes "leyes establecidas" y, en especial, los tres modelos típicos de los inductivistas [17]:

a) que el sol saldrá mañana y se pondrá cada veinticuatro horas (o aproximadamente cada 90.000 latidos del corazón),
b) que todos los hombres son mortales,
c) que el pan alimenta.

De hecho, en los tres casos, estas leyes establecidas fueron refutadas en el sentido en que originalmente se propusieron.

a) La primera la refutó Piteas de Marsella cuando descubrió "el mar helado y el sol de media noche". El que (*a*) significaba "vayas donde vayas el sol saldrá y se pondrá cada veinticuatro horas" queda de relieve por la manifiesta desconfianza con que fue recibida su narración, ya que se convirtió en el paradigma de todos los cuentos de aventuras.

b) La segunda también se vio refutada, aunque no de una manera tan obvia. El predicado "mortal" es una mala traducción del griego: θνητός significa "abocado a morir" o "sujeto a la muerte" más bien que simplemente "mortal". Precisamente (*b*) forma parte de la teoría aristotélica, según la cual toda criatura engendrada está abocada a degenerar y morir tras un período cuya extensión, aunque forma parte de la esencia de la criatura, puede variar un tanto de acuerdo con circunstancias accidentales. No obstante, esta teoría se vio refutada por el descubrimiento de que las bacterias no están abocadas a la muerte, ya que multiplicarse por fisión no es morir. También se vió refutada más tarde al comprobarse que la materia viva no está en general condenada a la degeneración, aunque parezca que con medios suficientemente drásticos es posible matar cualquiera de sus formas. (Las células cancerosas, por ejemplo, pueden continuar viviendo.)

c) La tercera —favorita de Hume— se vio refutada cuando la gente que comía su pan cotidiano murió de ergotismo en una catástrofe ocurrida en Francia no ha mucho. Naturalmente, lo que (*c*) significa propiamente es que el pan bien cocido, hecho de harina obtenida como es debido a partir del trigo o del maíz, plantado y recogido según la vieja costumbre, es más

---

[17] En el capítulo 2 (págs. 97 y sigs. y nota 58) también aparecen estos ejemplos que tantas veces he puesto en mis conferencias. Pido disculpas por la repetición, pero escribí ambos capítulos independientemente y considero que debo dejarlos tal como están.

bien alimento que veneno. Pero, con todo, aquellas personas se *envenenaron.*

Así pues, la respuesta negativa de Hume a H_L y mi respuesta negativa a L_1 no son simples actitudes filosóficas sofisticadas, como parecen dar a entender Ryle y la teoría del conocimiento del sentido común, sino que se basan en realidades muy prácticas. El profesor Strawson, con un tono optimista similar al del profesor Ryle, escribe: "Si ...hay un problema de la inducción y... Hume fue quien lo planteó, hay que añadir que fue él quien lo resolvió." Alude con esto a la respuesta positiva que dio Hume a H_{PS}, respuesta que Strawson parece aceptar describiéndola como sigue: "La Naturaleza nos fuerza... a aceptar los cánones básicos (de la inducción). La razón es, y debe ser, la esclava de las pasiones" [18]. (Hume había dicho: "Debe ser tan sólo"). Nunca había visto nada que ilustrase tan bien la cita de *A History of Western Philosophy* (pág. 699) de Bertrand Russell que he elegido como lema de la presente discusión.

Está claro que la "inducción" —en el sentido de una respuesta positiva a H_L o L_1— es *inductivamente inválida* e incluso paradógica, ya que una respuesta positiva a L_1, implica que nuestra explicación del mundo es aproximadamente verdadera (con lo cual estoy de acuerdo, a pesar de mi respuesta negativa a L_1). De ahí se sigue que somos un tipo de animales muy inteligentes, situados precariamente en un medio que difiere de cualquier otro lugar del universo: animales que se esfuerzan valerosamente en descubrir por cualquier medio las verdaderas regularidades que rigen el universo y, por tanto, nuestro medio. Mas es obvio que, sea cual sea el método que usemos, la probabilidad de encontrar regularidades verdaderas es escasa, y nuestras teorías estarán afectadas por errores en los que no nos impedirá incurrir ningún enigmático "canon de inducción", sea o no básico. Esto es precisamente lo que dice mi respuesta negativa a L_1. Por tanto, la respuesta positiva debe ser falsa, puesto que entraña su propia negación.

Quien quiera sacar la moraleja de esta historia deberá concluir que la razón crítica es mejor que la pasión, sobre todo en cuestiones de tipo lógico. Con todo, estoy dispuesto a admitir que no se puede hacer nada sin una pequeña dosis de pasión.

(4) L_2 no es más que una generalización de L_1, y L_3 no es sino una formulación alternativa de L_2.

(5) Mi respuesta a L_2 y L_3 suministra una respuesta clara a las preguntas de Russel. Puedo decir perfectamente que al menos algunos de los desvaríos de los lunáticos se pueden refutar mediante la experiencia, es decir, mediante enunciados contrastadores. (Otros pueden no ser contrastables, distinguiéndose así de las teorías científicas, lo cual plantea el problema de la demarcación [19].)

---

[18] Véase *Philosophical Studies*, 9, 1958, núms. 1-2, págs. 20 y sigs.; cf. Hume *Treatise*, pág. 415 [pág. 208 del volumen II de la traducción].

[19] Por "el problema de la demarcación" entiendo el problema de dar con un criterio mediante el cual podamos distinguir los enunciados de la ciencia de los enun-

(6) Lo que es más importante, mi respuesta a L_2, como subrayé en mi primer trabajo sobre el problema de la inducción, está de acuerdo con la siguiente forma —un tanto débil— del *principio del empirismo*: *Sólo la experiencia puede ayudarnos a decidir sobre la verdad o falsedad de los enunciados fácticos*. Esto es así, ya que, en vista de L_1 y de mi respuesta, podemos determinar al menos la falsedad de las teorías, lo cual se puede hacer también en vista de la respuesta a L_2.

(7) De un modo similar, mi solución no entra en conflicto con los métodos de la ciencia, sino que nos conduce a los rudimentos de una metodología crítica.

(8) Mi solución no sólo ilumina poderosamente el problema psicológico de la inducción (véase más adelante la sección 11), sino que también dilucida las formulaciones tradicionales del problema y las razones de su debilidad. (Véase las secciones 12 y 13, más adelante).

(9) Mis formulaciones y soluciones de L_1, L_2 y L_3 caen de lleno en el campo de la *lógica deductiva*. Muestro que, al generalizar el problema de Hume, podemos añadir L_2 y L_3, lo cual nos permite formular una respuesta algo más positiva que la que se puede dar a L_1. La causa estriba en la asimetría que existe, desde el punto de vista de la lógica deductiva, entre verificar y falsar mediante la experiencia. Esto conduce a la distinción exclusivamente lógica entre hipótesis que han sido refutadas y otras que no lo han sido y a la preferencia por estas últimas, aunque sólo sea desde un punto de vista teórico según el cual son *objetos teóricamente más interesantes para ulteriores contrastaciones*.

7. La preferencia entre teorías y búsqueda de la verdad

Hemos visto que la respuesta negativa a L_1 significa que todas nuestras teorías no son más que suposiciones, conjeturas o hipótesis. Una vez aceptado plenamente este resultado puramente lógico, surge la cuestión de si puede haber argumentos puramente racionales (que pueden ser empíricos) para preferir unas conjeturas o hipótesis a otras.

Hay varias maneras de considerar esta cuestión. Distinguiré el punto de vista del teórico —el que busca la verdad, en especial teorías explicativas verdaderas— del punto de vista del hombre de acción. Es decir, distinguiré la *preferencia teórica y la preferencia pragmática*. Tanto en ésta como en la próxima sección, me ocuparé exclusivamente de la preferencia teórica y la busca de la verdad. La preferencia pragmática y el problema de la "fiabilidad" se discutirán en la sección siguiente.

---

ciados no empíricos. Mi solución consiste en decir que un enunciado es empírico si hay conjunciones (finitas) de enunciados empíricos singulares ("enunciados básicos" o "enunciados contrastadores") que lo contradigan. De este "principio de demarcación" se sigue la consecuencia de que un enunciado puramente existencial aislado (por ejemplo, "En algún momento hay en alguna parte del mundo una serpiente marina") no es un enunciado empírico, aunque naturalmente, puede contribuir a plantear nuestra situación problemática empírica.

Supondré que el teórico se preocupa sólo por la verdad y, específicamente, por encontrar teorías verdaderas. Mas, una vez asimilado el hecho de que nunca podemos justificar empíricamente —es decir, por medio de enunciados contrastadores— la tesis de que una teoría es verdadera y que, por tanto, nos enfrentamos a lo sumo con el problema de preferir tentativamente unas suposiciones a otras, entonces, desde el punto de vista de quien busca teorías verdaderas, debemos atender a las preguntas: *¿Qué criterios de preferencia hemos de adoptar? ¿Hay teorías "mejores" que otras?*

Estas preguntas dan pie a las siguientes consideraciones.

(1) Está claro que el problema de la preferencia surgirá fundamentalmente —tal vez exclusivamente— en relación con conjuntos de *teorías rivales;* es decir, teorías que se ofrecen como soluciones a los mismos problemas. (Véase también el punto (8), más abajo.)

(2) El teórico que se interesa por la verdad, debe interesarse también por la falsedad, pues descubrir que un enunciado es falso equivale a descubrir que su negación es verdadera. Por tanto, la refutación de una teoría posee siempre un interés teórico, pero la negación de una teoría explicativa no es, a su vez, una teoría explicativa (ni posee, por regla general, el "carácter empírico" del enunciado contrastador del cual se deriva). Por muy interesante que sea, no satisface el interés que tiene el teórico por encontrar teorías explicativas verdaderas.

(3) Si el teórico persigue este fin, entonces descubrir dónde falla una teoría, además de suministrar una información teóricamente interesante, plantea un nuevo *problema* importante para una nueva teoría explicativa. Toda teoría nueva no sólo tiene que tener éxito donde lo tenía la teoría anterior refutada, sino que debe de tener éxito también donde ésta fallaba; es decir, en el punto en que fue refutada. Si la nueva teoría tiene éxito en ambos casos, será en cierta medida más afortunada y, por tanto, "mejor" que la vieja.

(4) Además, suponiendo que en el momento t la nueva teoría no se vea refutada por una nueva contrastación, también en otro sentido será "mejor" que la teoría refutada. En efecto, no sólo explicará todo lo que explicaba la teoría refutada más otras cosas, sino que además será considerada como posiblemente verdadera, ya que en el momento t no se ha mostrado su falsedad.

(5) No obstante, el teórico apreciará tal teoría, no sólo por su éxito y su posible verdad, sino también por su posible falsedad. Es interesante como objeto de subsiguientes contrastaciones, es decir, de nuevos intentos de refutación que, de tener éxito, no sólo establecerán una nueva negación de la teoría, sino también un nuevo problema teórico para la próxima teoría.

Podemos resumir los puntos (1)-(5) del modo siguiente:

Por diversas razones, el teórico se interesa por las teorías no refutadas, sobre todo porque algunas de ellas *pueden* ser verdaderas. Preferirá una teoría no refutada a una refutada, con tal de que explique los éxitos y fallos de la teoría refutada.

(6) Pero la teoría nueva puede ser falsa, como toda teoría no refutada. Por eso el teórico intentará por todos los medios detectar cualquier teoría falsa en el conjunto de las competidoras no refutadas; intentará "cazarla". O sea, intentará ingeniar circunstancias o situaciones en las cuales es probable que falle una teoría dada que no está refutada, si es que es falsa. Por tanto, se esforzará en construir contrastaciones *rigurosas* y situaciones *cruciales*, lo que puede entrañar la construcción de una ley falsadora, es decir, una ley que pueda ser quizá de un nivel de universalidad tan bajo que no sea capaz de explicar el éxito de la teoría a contrastar, aunque sugiera, no obstante, un *experimento crucial* que refute, según el resultado, o la teoría a contrastar o la teoría falsadora.

(7) Con este sistema de eliminación podemos dar con una teoría verdadera. Mas a pesar de que sea verdadera, este método no puede en ningún caso *establecer* su verdad, ya que el número de teorías verdaderas *posible* sigue siendo infinito en cualquier momento y tras cualquier número de contrastaciones cruciales. (Es otro modo de enunciar el resultado negativo de Hume.) El número de teorías efectivamente propuestas será obviamente finito, pudiendo ocurrir perfectamente que las refutemos todas y que no podamos inventar una nueva.

Por otra parte, *entre las teorías efectivamente propuestas* puede haber más de una sin refutar en un momento t, con lo que no sabremos cuál debemos preferir. Mas si en un momento t hay una pluralidad de teorías que siguen compitiendo de este modo, el teórico continuará con el intento de descubrir cómo diseñar experimentos cruciales entre ellas; es decir, experimentos que puedan falsar y eliminar consiguientemente algunas de las teorías rivales.

(8) El procedimiento descrito puede conducir a un conjunto de teorías que "compitan" en el sentido de ofrecer soluciones al menos a *algunos* problemas comunes, aunque cada una de ellas ofrezca por su parte solución a diversos problemas que no comparte con las otras. Aunque exigimos que una nueva teoría resuelva los problemas que resolvía su precesora *más* los que no resolvía, siempre puede ocurrir, como es natural, que se propongan dos o más teorías rivales nuevas, cada una de las cuales satisfaga estas exigencias y además resuelva algunos problemas que las otras no resuelven.

(9) El teórico puede estar especialmente interesado, en un momento t, en descubrir la teoría más contrastable para someterla a nuevas contrastaciones. Ya he mostrado que ésta ha de ser al mismo tiempo la que posea el mayor contenido informativo y el mayor poder explicativo. Será la teoría que más valga la pena someter a nuevas contrastaciones; resumiendo, "*la mejor*" de las teorías que compiten en un momento t. Si pasa las contrastaciones, será también *la mejor contrastada* de todas las teorías consideradas hasta el momento, incluyendo todas sus predecesoras.

(10) En lo dicho hasta aquí sobre "*la mejor teoría*" se ha supuesto que una buena teoría no es *ad hoc*. La idea de *caracter-ad-hoc* y su opuesta,

que tal vez se pueda denominar "audacia", son importantísimas. Las explicaciones *ad hoc* son las que no son contrastables independientemente; esto es, independientemente del efecto a explicar. Se formulan para un problema concreto, por lo que tienen escaso interés teórico.

Ya he discutido en otros lugares el problema de los grados de independencia de las contrastaciones [20]; es un problema interesante conectado con los de simplicidad y profundidad. He subrayado [21], desde entonces, la necesidad de referirlo y hacerlo depender del *problema de la explicación*, del que nos ocupamos ahora, y de las situaciones problemáticas discutidas, ya que todas estas ideas son importantes para los grados de "bondad" de las teorías rivales. Además, el grado de audacia de una teoría también depende de las relaciones que mantiene con sus predecesoras.

Creo que lo más interesante es que he conseguido dar un criterio objetivo para grados muy altos de audacia o *caracter-no-ad-hoc*. Ocurre que la nueva teoría, aunque debe explicar lo que explicaba la antigua, la *corrige* hasta el punto de que la *contradice* efectivamente: contiene la vieja teoría, *aunque sólo como aproximación*. En este sentido he señalado que la teoría de Newton contradice las teorías de Kepler y de Galileo —*si bien las explica* por el hecho de contenerlas como aproximaciones— y que la de Einstein contradice a la de Newton, que explica de un modo similar al contenerla como una aproximación.

(11) El método descrito puede denominarse *método crítico*. Es un método de ensayo y supresión de errores, de proponer teorías y someterlas a las contrastaciones más rigurosas que podamos diseñar. Si, mediante suposiciones limitadoras, sólo consideramos posibles un número finito de teorías rivales, el método puede llevarnos a señalar la teoría verdadera por eliminación de las competidoras. Normalmente —es decir, cuando el número de teorías posible es infinito— ni éste ni cualquier otro método pueden asegurar qué teoría es la verdadera. Aunque no concluyente, el método sigue siendo *aplicable*.

(12) El enriquecimiento de los problemas mediante la refutación de teorías falsas y las exigencias formuladas en (3), aseguran que la predecesora de cada una de las nuevas teorías sea —desde el punto de vista de ésta— una aproximación a esta nueva teoría. Naturalmente, nada asegura que hayamos de encontrar para cada teoría falsada una sucesora "mejor" o una aproximación mejor que satisfaga estas exigencias. *Nada asegura que podamos progresar hacia teorías mejores.*

(13) Aquí hay que añadir dos cosas. En primer lugar, lo dicho hasta ahora pertenece, como si dijéramos, a la *lógica puramente deductiva*, la lógica en la que se planteó L_1, L_2 y L_3. Sin embargo, al intentar su aplica-

[20] Véase especialmente "Naturgesetze und theoretische Systeme", en *Gesetz und Wirklichkeit*, ed. S. Moser, Insbruck 1949, págs 43 y sigs., así como "The Aim of Science", *Ratio, I*, 1957, recogidos ahora en el Apéndice y capítulo 5 respectivamente de este volumen.

[21] Véase *C. & R.*, pág. 241 [traducción castellana, pág. 279].

ción a las situaciones prácticas que surgen en la ciencia, chocamos con problemas de distinta índole. Por ejemplo, las relaciones entre enunciados contrastadores y teorías puede que no sean tan precisas como aquí suponemos; incluso es posible criticar los enunciados contrastadores mismos. Se trata del tipo de problemas que siempre surgen cuando deseamos *aplicar* la lógica pura a situaciones reales. En relación con la ciencia, conduce a lo que he llamado *reglas metodológicas,* reglas de discusión crítica.

En segundo lugar, puede considerarse que estas reglas están sometidas al *objetivo* general *de la discusión racional que consiste en acercarse a la verdad lo más posible.*

8. Corroboración: los méritos de la improbabilidad

(1) Mi teoría de la preferencia nada tiene que ver con la preferencia por las hipótesis "más probables". Por el contrario, he mostrado que la contrastabilidad de una teoría aumenta y disminuye con su *contenido informativo y,* por tanto, con su improbabilidad (en el sentido del cálculo de probabilidades). Así, lo más frecuente será que la hipótesis "mejor" o "preferible" sea la *más improbable.* (Con todo es un error decir, como hace John C. Harsanyi, que he propuesto un "criterio de improbabilidad para la elección de hipótesis científicas" [22]: no sólo no tengo un "criterio" general, sino que muy a menudo sucede que soy incapaz de preferir la hipótesis lógicamente "mejor" o más improbable, puesto que alguien ha conseguido refutarla experimentalmente. Como es natural, muchos han considerado que este resultado es molesto, pero mis argumentos fundamentales son sencillísimos (contenido = improbabilidad), y han sido recientemente aceptados incluso por algunos defensores del inductivismo y de la teoría probabilística de la inducción, como Carnap [23].

(2) Introduje originalmente la idea de *corroboración o "grado de corrobaración"* al objeto de mostrar claramente el carácter absurdo de toda teoría probabilística de la preferencia (y, por tanto, de toda teoría probabilística de la inducción).

Por grado de corroboración de una teoría entiendo un informe conciso que evalúe el estado (en un cierto momento *t*) de la discusión crítica de una teoría respecto al modo en que resuelve sus problemas, su grado de contrastabilidad, el rigor de las contrastaciones a que ha sido sometida y cómo ha salido de ellas. La corroboración (o grado de corroboración) es, por tanto, un *informe* evaluativo *de su rendimiento pasado.* Es esencialmente comparativo, como la preferencia: en general sólo puede decirse que una teoría *A* posee un grado de corroboración más elevado (o más bajo) que su rival *B*

[22] Véase John C. Harsanyi, "Popper's Improbability Criterion for the Choice of Scientific Hypotheses", *Philosphy,* 35, 1960, págs. 332-40. Véase también la nota de la página 218 de *C. & R.* [pág. 254 de la versión castellana].

[23] Véase Rudolf Carnap, "Probability and Content Mesure" en P. K. Feyerabend y Grover Maxwell (eds.) *Mind, Matter and Method,* Essays in Honour of Herbert Feigl, Univ. of Minnesota Press, Minneapolis 1966, págs. 248-60.

—a la luz de la discusión crítica que utiliza las contrastaciones— *hasta un momento dado, t.* Al ser un informe del rendimiento pasado, alude únicamente a un situación que puede llevar a preferir unas teorías a otras, *pero no dice nada de su rendimiento futuro ni de su "fiabilidad"*. (Por descontado, en nada afecta a esto el que alguien consiga mostrar que, en casos muy especiales, se pueda dar una interpretación numérica al grado de corroboración de mis fórmulas o a las de cualquier otro [24].)

El propósito fundamental de las *fórmulas* que propuse para definir el grado de corroboración era mostrar que en muchos casos es preferible la hipótesis más *improbable* (en el sentido del cálculo de probabilidades), señalando claramente en qué casos se puede decir esto y en cuáles no. Puedo mostrar así que *la preferibilidad no puede ser una probabilidad en el sentido del cálculo de probabilidades.* Naturaluente, se puede *llamar* preferible a la teoría más "probable": *no importan las palabras* con tal de que no nos dejemos engañar por ellas.

Resumiendo: A veces podemos decir de dos teorías rivales, *A* y *B*, que *A* es preferible, o está mejor corroborada, que la teoría *B*, a la luz del estado de la discusión crítica en el momento *t* y de la evidencia empírica (enunciados contrastadores) que tengamos a mano para la discusión.

Obviamente, el grado de corroboración en el momento *t* (que constituye un enunciado acerca de la preferibilidad en el momento *t*) no dice nada sobre el futuro —por ejemplo, sobre el grado de corroboración en un momento posterior a *t*. Exactamente, no es más que un informe sobre el estado, en el momento *t*, de la discusión acerca de la preferibilidad lógica y empírica de las teorías rivales.

(3) Debo subrayar esto porque el siguiente pasaje de mi libro *Logic of Scientific Discovery* ha sido interpretado —más bien mal interpretado— en el sentido de que yo empleaba la corroboración como índice del *rendimiento* futuro de una teoría: "en lugar de discutir la 'probabilidad' de una hipótesis deberíamos tratar de averiguar qué contrastaciones, qué pruebas ha soportado; esto es, tendríamos que intentar la averiguación de hasta qué punto ha sido capaz de demostrar que es apta para sobrevivir —y ello por haber salido indemne de las contrastaciones. En resumen, deberíamos disponernos a averiguar en qué medida está 'corroborada'" [25].

Algunas personas piensan [26] que la expresión "demostrar que es apta para sobrevivir" muestra que hablo aquí de la aptitud para sobrevivir en el *futuro*, para salir indemne de contrastaciones futuras. Si he confundido a alguien, lo siento. Lo único que puedo decir es que no he sido yo quien ha sacado de quicio la metáfora darwinista. Nadie supone que una especie

---

[24] Me parece que el profesor Lakatos sospecha que la aplicación efectiva de números a mi grado de corroboración, de ser posible, haría inductivista mi teoría, en el sentido de una teoría probabilista de la inducción. No veo por qué razón habría de ser así. Cf. págs. 410-12 de *The Problem of Inductive Logic*, I. Lakatos y A. Musgrave (eds.), North Holland, Amsterdam, 1968. (Añadido en las pruebas: me alegro de saber que he malinterpretado el pasaje aludido.)

[25] *L. Sc. D.*, pág. 251 [traducción citada, pág. 234].

[26] Véase *Mind*, New series, *69*, 1960, pág. 100.

que ha sobrevivido en el pasado vaya a sobrevivir en el futuro: todas las especies que no han logrado sobrevivir en un momento determinado *t* han sobrevivido hasta el momento *t*. Sería absurdo sugerir que la supervivencia darwinista implica de algún modo la expectativa de que continuará sobreviviendo toda especie que haya sobrevivido hasta este momento. (¿Quién osará decir que hay muchas esperanzas de que nuestra propia especie sobreviva?)

(4) Tal vez sea útil añadir aquí algo sobre el grado de corroboración de un enunciado *s* que pertenece o se deriva lógicamente de una teoría *T*, aunque sea lógicamente mucho más débil que ella.

Tal enunciado *s* tendrá menos contenido informativo que la teoría *T*, lo que significa que *s*, y el sistema deductivo *S* de todos aquellos enunciados que se siguen de *s*, será menos contrastable y corroborable que *T*. Pero si *T* ha sido bien contrastada, entonces podemos decir que su alto grado de corroboración se aplica a todos los enunciados implicados por ella y, por tanto, a *s* y *S*, aunque *s* nunca pudiese alcanzar por sí mismo un grado de corroboración tan alto, debido a su baja corroborabilidad.

Esta regla puede sostenerse considerando simplemente que el grado de corroboración es un medio de enunciar una *preferencia por lo que respecta a la verdad*. Así pues, si preferimos *T* por lo que respecta a su pretensión de verdad, entonces debemos preferirla junto con todas sus consecuencias, ya que si *T* es verdadera, deben serlo también todas sus consecuencias aunque separadamente puedan corroborarse peor.

Afirmo, por tanto, que con la corroboración de la teoría de Newton y la descripción de la tierra como un planeta en rotación, el grado de corroboración del enunciado *s*, "El sol sale en Roma cada venticuatro horas", ha aumentado considerablemente. La razón estriba en que *s*, por sí mismo, no puede contrastarse muy bien. Ahora bien, puesto que la teoría de Newton y la de la rotación terrestre pueden contrastarse bien, entonces, si son verdaderas, *s* también lo será.

Un enunciado *s*, derivable de una teoría *T* bien contrastada, tiene el mismo grado de corroboración que *T*, *en tanto en cuanto se considera* parte de *T*; y si *s* no es derivable de *T*, sino de la conjunción de dos teorías, T_1 y T_2, en cuanto parte de dos teorías, tendrá el mismo grado de corroboración que la peor contrastada de ambas. Sin embargo, en sí mismo *s* puede tener un grado de corroboración muy bajo.

(5) La diferencia fundamental entre mi enfoque y el enfoque que he denominado hace tiempo "inductivista" consiste en que yo pongo el acento en los *argumentos negativos*, tales como casos negativos o contra-ejemplos, refutaciones e intentos de refutación —brevemente, crítica—, mientras que el inductivista pone el acento en los "*casos positivos* de los que saca *inferencias* no-demostrativas" [27] que pretende que garanticen la "*fiabilidad*" de las conclusiones de estas inferencias. Para mí, lo único que puede ser "*po-*

[27] C. G. Hempel, "Recent Problems of Induction", en R. G. Colodny (ed.), *Mind and Cosmos*, Pittsburgh Univ. Press, 1966, pág. 112.

sitivo" en el conocimiento científico *sólo* es positivo en tanto en cuanto ciertas teorías son, en un momento dado, preferidas a otras a la luz de nuestra discusión *crítica* consistente en intentos de refutación que incluye contrastaciones empíricas. Por tanto, incluso lo que puede ser considerado "positivo", lo es *sólo* por respecto a *métodos negativos.*

Este enfoque negativo clarifica muchas cosas; por ejemplo, las dificultades que se encuentran al explicar satisfactoriamente lo que constituye un "caso positivo" o un "caso favorable" de una ley.

9. Preferencia pragmática

He discutido hasta aquí por qué el teórico ha de preferir —si es que tiene alguna preferencia— la teoría "mejor", es decir, la más contrastable y la mejor contrastada. Naturalmente, el teórico puede no tener ninguna preferencia: puede sentirse desanimado por la solución "escéptica", mía y de Hume, a los problemas H_L y L_1; puede alegar que si no somos capaces de *asegurar* el descubrimiento de cuál de las teorías rivales es la verdadera, no le interesa para nada semejante método —ni siquiera si el método hace razonablemente cierto que *si* hubiese una teoría verdadera entre las propuestas, estaría entre las supervivientes, preferidas o corroboradas. Sin embargo, puede ser que un teórico "puro" más entusiasta o más curioso se anime con nuestro análisis a proponer una y otra vez nuevas teorías rivales, confiando en que tal vez una de ellas sea verdadera— aunque nunca seamos capaces de asegurar que alguna de ellas sea verdadera.

El teórico puro tiene, pues, ante sí más de una vía de acción: eligirá un método como el de ensayo y supresión de errores sólo si su curiosidad supera su frustración por la inevitable incertidumbre y deficiencia de todos nuestros intentos.

El caso es distinto si lo consideramos como hombre de acción práctica, ya que éste ha de *elegir* siempre entre algunas alternativas más o menos concretas, puesto que incluso *la inacción es un tipo de acción.*

Mas toda acción presupone un conjunto de expectativas; es decir, de teorías sobre el mundo. ¿Qué teoría elegirá el hombre de acción? ¿Se puede hablar de una *elección racional?*

Esto nos lleva a los *problemas pragmáticos de la inducción*:

Pr_1 ¿De qué teoría hemos de fiarnos desde un punto de vista racional, para la actividad práctica?

Pr_2 ¿Qué teoría hemos de preferir, desde un punto de vista racional, para la actividad práctica?

Mi respuesta a Pr_1 es la siguiente: Desde un punto de vista racional no podemos "fiarnos" de ninguna teoría, ya que no se ha demostrado, ni se puede demostrar, que una teoría sea verdadera.

Mi respuesta a Pr_2 es: sin embargo, debemos *preferir* la teoría mejor contrastada como base de acción.

En otras palabras, no hay "confianza absoluta", pero puesto que *hemos* de elegir, será "racional" elegir la teoría mejor contrastada. Será "racional" en el sentido más obvio de la palabra que conozco: la teoría mejor contrastada es la que a la luz de la *discusión crítica* parece mejor, por el momento, y no conozco nada más "racional" que una discusión crítica bien llevada.

Evidentemente, al elegir la teoría mejor contrastada como base de acción, "confiamos" en ella en algún sentido de la palabra. Por tanto, en cierto sentido, incluso se la puede describir como la *más* "fiable" de las teorías utliizables. Sin embargo, esto no quiere decir que sea "fiable", al menos en el sentido de que siempre procederemos prudentemente si, incluso en la acción práctica, prevemos la posibilidad de que nuestras expectativas fallen en algo.

No es sólo esta preocupación trivial lo que debemos derivar de nuestra respuesta negativa a L_1 y P_1. Por el contrario, para la plena comprensión del problema —especialmente de lo que he denominado el problema tradicional— es de la mayor importancia percatarse de que, a pesar de la "racionalidad" de elegir la teoría mejor contrastada como base de acción, esta elección *no* es "racional" en el sentido de apoyarse en *buenas razones* para esperar que resulte una elección afortunada en la práctica: *no puede haber buenas razones* en este sentido, que es en lo que consiste precisamente el resultado de Hume. (En esto coinciden plenamente nuestras respuestas a H_L, L_1 y Pr_1.) Por el contrario, aunque muchas teorías físicas fuesen verdaderas, es perfectamente posible que el mundo, tal como lo conocemos con todas su regularidades pragmáticamente relevantes, se desintegre completamente el segundo que viene. Hoy día, esto debería ser obvio para todos; yo ya lo dije [28] antes de Hiroshima: es infinitamente posible que ocurra un desastre local, total o parcial.

No obstante, desde un punto de vista pragmático, es obvio que no vale la pena preocuparse por la mayor parte de estas posibilidades, ya que nada podemos *hacer*: están más allá del campo de acción. (Naturalmente, no incluyo la guerra atómica entre los desastres que están más allá del campo de acción humana, aunque así lo crea la mayor parte de la gente porque es tan incontrolable para la mayoría como una acción divina.)

Todo esto valdría igual, aún cuando pudiésemos estar seguros de la verdad de nuestras teorías físicas y biológicas. Pero no lo estamos, sino que, por el contrario, poseemos razones para sospechar incluso de las mejores, lo cual añade, naturalmente, nuevos infinitos a las infinitas posibilidades de desastre.

Este tipo de consideraciones son las que hacen tan importantes mi respuesta negativa y la de Hume, ya que ahora podemos ver con toda claridad por qué hemos de guardarnos de que nuestra teoría del conocimiento pruebe demasiado. Para decirlo con más exactitud, *ninguna teoría del conocimiento*

[28] Véase *L. d F.*, sección 79 (*L. Sc. D.*, págs. 253 y sigs. [versión castellana, página 236].)

puede pretender explicar por qué tienen éxito nuestros intentos de explicar las cosas.

Aunque supongamos que hemos tenido éxito —que nuestras teorías físicas son verdederas— nuestra cosmología puede enseñarnos cuán infinitamente improbable es este éxito: nuestras teorías nos muestran que el mundo está casi completamente vacío y que el espacio vacío está lleno de radiaciones caóticas. Por otro lado, la mayoría de las partes no vacías están ocupadas por polvo caótico, por gases o por estrellas a altísimas temperaturas —en condiciones que parecen impedir la aplicación de cualquier método para adquirir localmente un conocimiento físico.

Resumiendo, hay muchos mundos actuales y posibles en los que fracasará la búsqueda de regularidades y conocimiento. Incluso en el mundo, tal y como lo conocemos por las ciencias, la presencia de condiciones en las que pueda surgir —y prosperar— la vida y la búsqueda del conocimiento parece ser casi infinitamente improbable. Da la impresión, además, de que, aunque apareciesen tales condiciones en algún momento, estarían abocadas una vez más a la desaparición tras un período muy breve, cosmológicamente hablando.

10. Bases de mi replanteamiento del problema psicológico de la inducción de Hume

Históricamente hablando, descubrí la nueva solución al problema psicológico de la inducción de Hume antes que la solución al problema lógico: fue aquí donde por primera vez noté que la inducción —la formación de una creencia por repetición— es un mito. Primero en animales y niños y luego en adultos observé la *necesidad* inmensamente poderosa *de regularidades* —la necesidad que les hace buscar regularidades, que les hace verlas incluso donde no las hay, que les hace aferrarse dogmáticamente a sus expectativas, y que los hace desgraciados si se derrumban ciertas regularidades supuestas, pudiendo llevarlos a la desesperación y al borde de la locura. Cuando Kant dijo que nuestro intelecto impone sus leyes a la naturaleza, estaba en lo cierto— salvo que no se percató de cuán a menudo fracasa: las regularidades que intentamos imponer son *a priori psicológicamente,* pero no hay la menor razón para suponer que sean *válidas a priori,* como pensaba Kant. La necesidad que lleva a intentar imponer tales regularidades al medio es claramente innata, basándose en impulsos o instintos. Existe una necesidad general de que el mundo se conforme a nuestras expectativas, aunque hay muchas otras necesidades más específicas como son la necesidad de respuesta social regular o la necesidad de aprender un lenguaje con reglas para los enunciados descriptivos (o de otro tipo). Esta situación me hizo concluir, en primer lugar, que las expectativas pueden surgir sin, o antes de, la repetición y, en segundo lugar, me condujo a un análisis lógico que mostraba que no podían surgir de otra manera, puesto que la repetición presupone la similitud, que a su vez presupone un punto de vista —una teoría o una expectativa.

Por tanto, resolví que *por razones lógicas* no podía ser verdadera la teoría inductiva de Hume sobre la formación de creencias, lo cual me hizo ver que había que transferir las consideraciones lógicas al terreno psicológico. Esto, a su vez, me llevó a la ulterior conjetura heurística de que, por regla general, lo que vale en lógica vale también en psicología, suponiendo que se transfiera adecuadamente. (Este principio heurístico es el que ahora denomino "principio de transferencia".) Imagino que fue en gran medida este resultado el que me hizo abandonar la psicología en favor de la lógica del descubrimiento.

Aparte de esto, me dio la sensación de que habría que considerar la psicología como una disciplina biológica, especialmente las *teorías* psicológicas *sobre la adquisición de conocimientos.*

Ahora bien, si transferimos a la psicología humana y animal este *método de preferencia* que deriva de nuestra solución a L_3, llegamos claramente al método de sobra conocido de ensayo y supresión de errores: los diversos ensayos corresponden a la formación de hipótesis rivales y la supresión de errores, a la supresión o refutación de teorías mediante contrastaciones.

Esto me condujo a la siguiente formulación: la diferencia fundamental entre Einstein y una ameba (tal como describe Jennings [29]) estriba en que Einstein *busca conscientemente la supresión de errores.* Intenta matar sus teorías, *criticándolas conscientemente,* razón por la cual trata de *formularlas* no con vaguedad, sino con precisión. Mas la ameba no puede ser crítica frente a sus expectativas o hipótesis, no puede *plantarles cara*: forman parte de ella. (Sólo se puede criticar el conocimiento objetivo; el subjetivo sólo se puede criticar cuando se hace objetivo, lo cual ocurre cuando *decimos* lo que pensamos o, mejor aún, cuando lo *escribimos* o lo hacemos *imprimir.*)

Está claro que el método de ensayo y supresión de errores se basa, en gran medida, en instintos innatos, algunos de los cuales están ligados con ese vago fenómeno que algunos filósofos llaman "creencia".

Acostumbro a vanagloriarme de no ser un filósofo de la creencia: me intereso primariamente por las ideas —las teorías— y considero comparativamente sin importancia el que alguien "crea" o no en ellas. Sospecho que el interés de los filósofos por la creencia es un resultado de esa filosofía equivocada que denomino "inductivismo". Hay teóricos del conocimiento que, partiendo de experiencias subjetivas, no logran distinguir el conocimiento objetivo del subjetivo, lo que les lleva a pensar que la creencia es el género y el conocimiento una de sus especies. (La diferencia específica vendría dada por la "justificación" o tal vez por un "criterio de verdad" como la claridad y distinción, la vivacidad [30] o la "razón suficiente".)

Por eso no creo, como tampoco E. M. Forster, en la creencia.

Sin embargo, hay otras razones más importantes para desconfiar de la

---

[29] H. S. Jennings, *The Behaviour of the Lower Organisms,* Columbia University, 1906.

[30] Véase Hume, *Treatise,* pág. 265. [Traducción citada, pág. 409 del tomo I.]

creencia. Estoy totalmente dispuesto a admitir que algunos estados psicológicos pueden denominarse "expectativas" y que hay matices, desde la vívida expectativa del perro a quien se va a sacar de paseo, hasta la casi inexistente del escolar que sabe, sin creerlo realmente, que llegará el día en que será un viejo, si es que vive lo suficiente. Pero es discutible que los filósofos utilicen la palabra "creencia" para describir estados psicológicos en este sentido. Parece que a lo que se refieren, más bien, no es a estados momentáneos, sino a lo que podemos llamar creencias "fijas", incluso aquellas innumerables expectativas inconscientes que constituyen nuestro horizonte de expectativas. Hay un gran trecho de éstas a las hipótesis formuladas y, por tanto, también a los enunciados del tipo "creo que...".

Ahora bien, la mayoría de estos enunciados *formulados* se pueden considerar críticamente; y los estados psicológicos que *resultan* de una consideración crítica me parecen muy distintos también de una expectativa inconsciente. Por tanto, incluso una creencia "fija" cambia cuando se la formula, así como después de su formulación. Si el resultado de su consideración crítica es la "aceptación", puede dejar de ser una aceptación fanática que amenaza con suprimir las dudas y escrúpulos para convertirse en esa aceptación provisional dispuesta a la reconsideración y revisión según nuevos datos y que incluso puede estar ligada a una activa búsqueda de refutaciones.

No creo que estas distinciones entre diversas "creencias" tengan algún interés para mi teoría objetivista del conocimiento; pero deberían interesar a quien se tome en serio el problema psicológico de la inducción —cosa que yo no hago.

11. Replanteamiento del problema psicológico de la inducción

Por las razones expuestas, no considero que el problema psicológico de la inducción forme parte de mi propia teoría (objetivista) del conocimiento. Con todo, creo que el principio de transferencia sugiere los siguientes problemas y respuestas.

Ps$_1$ Si consideramos críticamente una teoría desde el punto de vista de los elementos de juicio suficientes en su favor más bien que desde un punto de vista pragmático, ¿tenemos siempre un sentimiento de completa seguridad o certeza acerca de su verdad, incluso por lo que se refiere a las teorías mejor contrastadas, como que el sol sale todos los días?

Creo que la respuesta debe ser: No. Creo que el sentimiento de certeza —la creencia fuerte— que Hume intentaba explicar, es una creencia *pragmática*; algo estrechamente relacionado con la acción y la elección entre alternativas o, incluso, con la necesidad y expectativa de regularidades. Pero si suponemos que estamos en disposición de reflexionar sobre la evidencia y lo que nos permite afirmar, entonces hemos de admitir que, después de todo, el sol puede no salir mañana en Londres —por ejemplo—, porque puede explotar dentro de media hora, con lo que no habrá mañana. Natu-

ralmente, no hemos de tomar "seriamente" —es decir, pragmáticamente— esta posibilidad, puesto que no sugiere ninguna acción posible: no podemos hacer nada con ella.

Por tanto, debemos considerar nuestras creencias pragmáticas que pueden ser muy fuertes. Preguntamos:

Ps_2 ¿Esas "fuertes creencias pragmáticas" que todos mantenemos, como es la creencia en que habrá un mañana, son el resultado irracional de la repetición?

Mi respuesta es: No. La teoría de la repetición es, en todos los sentidos, insostenible. Estas creencias son en parte innatas, en parte modificaciones de creencias innatas que surgen del método de ensayo y supresión de errores. Mas este método es perfectamente "racional", ya que corresponde precisamente a ese método de preferencia cuya racionalidad hemos discutido Más explícitamente, una *creencia pragmática en los resultados de la ciencia* no es irracional, ya que nada hay más "racional" que el método de la discusión crítica que es el método de la ciencia. Aunque fuese irracional aceptar como cierto cualquiera de sus resultados, no hay nada "mejor" a la hora de actuar en la práctica: no hay otro método alternativo que pueda considerarse más racional.

12. El problema tradicional de la inducción y la invalidez de todos los principios o reglas de inducción

Vuelvo ahora a lo que he llamado el problema filosófico de la inducción.

Con ésto aludo a lo que ocurre si aceptamos el punto de vista del sentido común acerca de la inducción por repetición puesto en entredicho por Hume, sin tomar la crítica con la debida seriedad. Si después de todo, el propio Hume siguió siendo inductivista, no podemos esperar que los inductivistas criticados por él tomen en serio su crítica.

El esquema básico del problema tradicional se puede plantear de varias maneras; por ejemplo:

Tr_1 ¿Cómo se puede justificar la inducción (a pesar de Hume)?

Tr_2 ¿Cómo se puede justificar un principio de inducción (es decir, un principio que justifique la inducción y que no sea lógico)?

Tr_3 ¿Cómo se pede justificar un principio de inducción, del tipo "el futuro será como el pasado" o, tal vez, el llamado "principio de uniformidad de la naturaleza"?

Como indiqué brevemente en mi libro *Logik der Forschung*, considero que el problema kantiano, "¿Cómo pueden ser válidos *a priori* los enunciados sintéticos?", constituye un intento de generalizar Tr_1 o Tr_2. Por eso tengo a Russell por kantiano, ya que en algunas de sus fases intentó solucionar Tr_2 mediante una justificación *a priori*. En los *Problems of Philosophy*,

por ejemplo, Russell formuló Tr_2 como sigue "...¿qué clase de creencias generales serán suficientes, si fueran verdaderas, para justificar el juicio según el cual el sol saldrá mañana...?".

Para mí todos estos problemas están mal formulados. Como ocurre con las versiones probabilísticas del tipo del principio de inducción de Thomas Reid, "lo que ha de ser, será probablemente como lo que ha sido en circunstancias similares". Sus autores no han tomado con suficiente serenidad la crítica lógica de Hume y nunca toman en serio la posibilidad de que podamos, y debamos, arreglárnoslas sin la inducción por repetición, cosa que de hecho ocurre.

Me parece que todas las objeciones que conozco a mi teoría se formulan bajo el punto de vista de si ha resuelto o no el problema tradicional de la inducción —es decir, si he justificado o no la inferencia inductiva. Como es evidente que no es ese el caso, mis críticos deducen que no he conseguido resolver el problema de la inducción de Hume.

Hay que rechazar las formulaciones tradicionales del principio de inducción por diversas razones, especialmente por las enunciadas en la sección 9. Han de ser rechazadas porque suponen, no sólo que nuestra busca de conocimiento ha tenido éxito, sino también que hemos de poder explicar por qué.

No obstante, aun suponiendo (lo que comparto) que nuestra busca de conocimiento ha tenido éxito hasta ahora y que sabemos algo acerca del universo, este éxito es milagrosamente improbable y, por ende, inexplicable, ya que apelar a una serie ilimitada de accidentes improbables no constituye una explicación. (Supongo que lo único que podemos hacer es investigar la casi increíble historia evolutiva de estos accidentes desde la formación de los elementos hasta la de los organismos.)

Una vez hecho esto, se verá que es totalmente obvia, no sólo la tesis de Hume de que apelar a la probabilidad no puede cambiar la respuesta a H_L (y, por tanto, a L_1 y Pr_1), sino también la invalidez de cualquier "principio de inducción."

La idea de un principio de inducción es la de que hay un enunciado —que ha de considerarse como un principio metafísico válido *a priori*, probable o, tal vez, como una mera conjetura— que, de ser verdadero, suministrará *buenas razones para que confiemos en regularidades*. Si por "confianza" entendemos simplemente confianza pragmática en la racionalidad de nuestras *preferencias* teóricas, en el sentido de Pr_2, entonces será claro que no necesitamos ningún principio de inducción: no precisamos confiar en regularidades —es decir, en la verdad de las teorías— para justificar esta preferencia. Si, por el contrario, se alude a la "confianza" en el sentido de Pr_1, entonces tal principio de inducción será sencillamente falso. Es más, incluso será paradójico en el siguiente sentido: nos permitiría confiar en la ciencia, cuando ésta nos enseña hoy día que sólo en condiciones muy especiales e improbables pueden tener lugar situaciones que permitan observar regularidades o casos de regularidades. De hecho, la ciencia nos enseña que

tales condiciones difícilmente aparecen en alguna parte del universo y, si aparecen (digamos, en la tierra), será durante períodos muy cortos desde un punto de vista cosmológico.

Está claro que esta crítica no sólo se aplica a cualquier principio que justifique la inferencia inductiva basada en la repetición, sino también a cualquier principio que justifique la "confianza", en el sentido de Pr_1, en el método de ensayo y supresión de errores o en cualquier otro método posible.

13. Más allá de los problemas de inducción y demarcación

La solución al problema de la inducción se me ocurrió mucho después de resolver a mi entera satisfacción el problema de la demarcación (demarcación entre ciencia empírica y pseudo-ciencia, especialmente la metafísica).

Sólo tras solucionar el problema de la inducción consideré que el problema de la demarcación era objetivamente importante, ya que antes sospechaba que se reducía a suministrar una definición de la ciencia, lo cual me parecía de significación dudosa (tal vez por mi actitud negativa hacia las definiciones), aunque lo encontraba muy útil para clarificar mi posición frente a la ciencia y la pseudo-ciencia.

Vi que lo que había que eliminar era la *busca de justificaciones*, en el sentido de justificar la pretensión de verdad de una teoría. *Todas las teorías son hipótesis*, todas *pueden* ser rechazadas.

Sin embargo, no sugería, ni mucho menos, que hubiese que eliminar la busqueda de la verdad: nuestra discusión crítica de las teorías está presidida por la idea de encontrar una teoría explicativa verdadera (y poderosa). Además, *justificamos nuestras preferencias apelando a la idea de verdad* que desempeña el papel de idea reguladora. *Contrastamos* para encontrar la *verdad*, eliminando la falsedad. El que no podamos suministrar una justificación —o razones suficientes— de nuestras sospechas no quiere decir que no podamos haber rastreado la verdad; muchas de nuestras hipótesis pueden perfectamente ser verdaderas [31].

Cuando nos percatamos de que todo conocimiento es hipotético, rechazamos el "principio de razón suficiente", en el sentido de "que se puede dar una razón para toda verdad" (Leibniz) o en el sentido más fuerte que encontramos en Berkeley y Hume, quienes sugirieron que hay una razón suficiente para no creer si "no vemos razones [suficientes] para creer" [32].

Una vez que resolví el problema de la inducción y me di cuenta de

---

[31] Ni siquiera hace falta decirlo. Sin embargo la *Encyclopedia of Philosophy*, 1967, vol. *3*, pág. 37, me atribuye la tesis de que "la verdad no es más que una ilusión".

[32] Berkeley, *Three Dialogues Betwen Hylas and Philonous*, Segundo diálogo: "...es razón suficiente para no creer... el que yo vea que no hay razón suficiente para creer". [Hay traducción castellana de Vicente Viqueira, Madrid, Calpe, 1923, pág. 94.]

Para el caso de Hume véase *C. & R.*, pág. 21 [traducción castellana, pág. 30] (donde se cita el *Enquiry Concerning Human Understanding*, Sec. V, parte I).

sus estrechas relaciones con el problema de la demarcación, surgieron, uno tras otro, atropellándose, nuevos problemas interesantes y nuevas soluciones.

Ante todo, me di cuenta en seguida de que el problema de la demarcación y mi solución, tal como la he enunciado más arriba, eran un tanto formales y poco realistas: *siempre se pueden evitar las refutaciones empíricas;* siempre se puede "*inmunizar*" una teoría contra la crítica. (Considero que esta excelente expresión, debida a Hans Albert, debe sustituir a mis expresiones "estratagema convencionalista" y "giro convencionalista".)

Así fue como llegué a la idea de las *reglas metodológicas* y de la importancia fundamental del punto de vista crítico; es decir, de un punto de vista que impida la política de inmunizar nuestras teorías contra la refutación.

Al mismo tiempo, constaté también lo opuesto: el valor de una actitud *dogmática;* alguien tenía que defender la teoría de la crítica o, de lo contrario, sucumbiría con demasiada facilidad antes de poder contribuir al crecimiento científico.

El paso siguiente consistió en aplicar el punto de vista crítico a los enunciados contrastadores, la "base empírica": subrayé el carácter hipotético de toda observación y de todo enunciado observacional.

Esto me llevó a pensar que todo lenguaje está impregnado de teoría, lo cual entraña, obviamente, una revisión radical del empirismo. También me hizo considerar la actitud crítica como característica de la actitud racional; la significación de la función argumentativa (o crítica) del lenguaje; la idea de la lógica deductiva como órgano de la crítica y la importancia de la transmisión de la falsedad de la conclusión a las premisas (corolario de la transmisión de la verdad de las premisas a la conclusión). También hizo que me percatase de que sólo puede ser objetiva una teoría *formulada* (en oposición a una teoría creída), de la idea de que esta formulación u objetividad es lo que posibilita la crítica y, consecuentemente, de mi teoría de un "tercer mundo" (o "mundo 3", como prefiere llamarlo Sir John Eccles) [33].

Estos son sólo algunos de los muchos problemas que surgieron del nuevo punto de vista. Otros problemas son de carácter más técnico, como los relacionados con la teoría de la probabilidad, teniendo en cuenta el papel que desempeña en la teoría cuántica y la conexión entre mi teoría de la preferencia y la teoría darwinista de la selección natural.

---

[33] John F. Eccles, *Facing Reality*, Springer-Verlag, Berlín-Heidelberg-New York, 1970.

2. LAS DOS CARAS DEL SENTIDO COMUN: ARGUMENTOS EN PRO DEL REALISMO DEL SENTIDO COMUN Y EN CONTRA DE LA TEORIA DEL CONOCIMIENTO DEL SENTIDO COMUN*

1. Defensa de la filosofía

En los tiempos que corren es imprescindible pedir disculpas por ocuparse de la filosofía en cualquiera de sus formas. Exceptuando tal vez algunos marxistas, la mayor parte de los filósofos profesionales parecen haber perdido contacto con la realidad. Sin embargo: "los marxistas se han limitado a *interpretar* el maxismo de diversas maneras; lo que importa, no obstante, es *transformarlo*" [1].

En mi opinión, el mayor escándalo filosófico consiste en que, mientras a nuestro alrededor el mundo natural —y no sólo el mundo natural— se hunde, los filósofos continúan debatiendo, inteligentemente o no, el problema de si tal mundo existe. Se ven envueltos en el escolasticismo [2] y en enredos lingüísticos como, por ejemplo, si hay diferencias o no entre "ser" y "existir". (Como en el arte contemporáneo, no hay criterios para estos términos filosóficos.)

No hace falta decir que la extendida actitud anti-intelectual, tan acentuada entre los nacional-socialistas y que cada día se acentúa más entre los jóvenes desilusionados, especialmente estudiantes, es tan perniciosa como este tipo de escolasticismo y, si ello es posible, algo peor incluso que la verborrea pretenciosa y espúrea, aunque a veces brillante, de los filósofos y otros intelectuales. Pero es sólo algo peor, ya que la deslealtad de los

---

* Este largo ensayo, aún inédito, es una versión revisada y aumentada de una charla dada en mi seminario en 1970. Pretende ser una respuesta exhaustiva a las críticas que se han hecho a mis opiniones sobre la ciencia. Estoy en deuda con John Watkins que ha leído una primera versión de este ensayo y me señaló un serio error que, por fortuna, no resultó ser relevante para el argumento general. David Miller ha dedicado su tiempo con toda generosidad a leer el ensayo una y otra vez, salvándome no sólo de tres errores similares, sino también de innumerables oscuridades menores de contenido y estilo por lo que le estoy profundamente agradecido.

[1] Naturalmente, fue Marx quien dijo (en la undécima de las *Tesis sobre Feuerbach*): "Los filósofos se han limitado a *interpretar* el mundo de diversas maneras; sin embargo, lo que importa es *transformarlo*". La brillante y oportuna variación citada en el texto parece deberse a R. Hochhuth. (Mas no debo mencionar la brillantez de Hochhut sin repudiar enérgicamente su actitud equivocada hacia Winston Churchill.)

[2] Empleo el término "escolasticismo" para aludir a la tendencia a argumentar sin ningún problema serio, actitud que no era en absoluto general entre los escolásticos de la Edad Media.

intelectuales despierta el anti-intelectualismo como una reacción casi inevitable. Si se alimenta a los jóvenes con piedras en lugar de con pan, se rebelarán y tomarán a un panadero por alguien que tira piedras.

En estas circunstancias hay que disculparse por ser filósofo y especialmente por replantear (según pretendo) una trivialidad como el *realismo*, la tesis de la realidad del mundo. ¿Cuál es mi escusa? Hela aquí:

Seamos o no conscientes de ello, todos tenemos una filosofía propia que no vale gran cosa. Sin embargo, su impacto sobre nuestras acciones y vidas puede llegar a ser devastador, lo cual hace necesario tratar de mejorarla mediante la crítica. Es la única disculpa que puedo dar de la persistente existencia de la filosofía.

2. Punto de partida incierto: sentido común y crítica

La ciencia, la filosofía, el pensamiento racional deben surgir todos del sentido común. Sin embargo, el sentido común no es un punto de partida seguro: el término "sentido común" que aquí empleo es muy vago, porque denota algo vago y cambiante —los instintos u opiniones de la gente, muchas veces adecuados y verdaderos, pero muchas otras inadecuados o falsos.

¿Cómo es posible que una cosa tan vaga e insegura como el sentido común nos suministre un punto de partida? Mi respuesta es: porque no intentamos ni pretendemos construir (como, por ejemplo, Descarte, Spinoza, Locke, Berkeley o Kant) un sistema seguro sobre esos "fundamentos". Todas nuestras diversas suposiciones de sentido común —nuestro conocimiento básico de sentido común, podríamos decir— de las que podríamos arrancar pueden ser criticadas y puestas en entredicho en cualquier momento. Es muy frecuente que tales suposiciones sean criticadas con éxito y rechazadas (por ejemplo, la teoría de que la tierra es plana). En tal caso, el sentido común, o bien es modificado tras la corrección, o bien es superado y reemplazado por una teoría que, durante un período de tiempo más o menos largo, puede parecer a algunas personas un tanto "extravagante". Si la comprensión de la teoría exige una gran formación, puede ocurrir que nunca consiga ser asimilada por el sentido común. Incluso entonces hemos de exigir el intento de acercarse lo más posible al ideal: *Toda ciencia y toda filosofía son sentido común ilustrado.*

Comenzamos, pues, con un punto de partida vago y construimos sobre una base insegura. Con todo, podemos progresar: algunas veces, tras una crítica, podemos ver que nos hemos equivocado; podemos aprender de nuestros errores, si nos percatamos de haber cometido un error.

(Incidentalmente, intentaré mostrar más adelante que el sentido común ha llevado a confusiones especialmente en la teoría del conocimiento: es la teoría equivocada según la cual adquirimos conocimientos sobre el mundo abriendo los ojos y mirando o, en general, observando.)

Mi primera tesis es que partimos del sentido común, siendo la crítica nuestro gran instrumento de progreso.

Sin embargo, esta tesis plantea inmediatamente una dificultad. Se ha dicho que si queremos criticar una teoría, digamos T_1, sea o no de sentido común, precisamos otra teoría, T_2, que nos suministre las bases necesarias, el punto de partida o el transfondo para criticar a T_1. Sólo en el caso muy especial de que podamos mostrar que T_1 es inconsistente (caso denominado "crítica inmanente", en el que empleamos T_1 para mostrar que T_1 es falsa) podemos proceder de otro modo; es decir, mostrando que de T_1 se siguen consecuencias absurdas.

No considero válida esta crítica del método crítico. (Alega que toda crítica ha de ser o "inmanente" o "transcendente" y que en el caso de la crítica "transcendente" no procedemos de un modo crítico porque hemos de suponer dogmáticamente que T_2 es verdadera.) Lo que realmente ocurre es lo siguiente: si creemos que hemos de formular una crítica a T_1, que podemos suponer consistente, entonces, o bien mostramos que T_1 lleva a consecuencias inesperadas o indeseables (no importa demasiado que sean lógicamente inconsistentes), o bien mostramos que hay otra teoría rival T_2 que choca con T_1 y tratamos de mostrar que tiene algunas ventajas sobre T_1. Sólo precisamos esto: tan pronto como disponemos de teorías rivales, hay amplio campo para la discusión crítica o racional: exploraremos las consecuencias de las teorías e intentaremos especialmente descubrir sus puntos débiles, es decir, consecuencias que consideramos equivocadas. Este tipo de discusión crítica o racional puede llevarnos al fracaso más rotundo de una de las teorías, pero lo más normal es que nos ayude a poner en claro la debilidad de ambas, lo que nos incita a formular otra teoría.

El problema fundamental de la teoría del conocimiento es la clarificación e investigación de ese proceso mediante el cual aumentan o progresan, como aquí pretendemos, nuestras teorías.

3. Diferencias con otros enfoques

Tal vez lo dicho hasta aquí parezca completamente trivial. Por ello, intentaré compararlo muy brevemente con otros enfoques.

Quizá fue Descartes el primero en afirmar que todo depende de la seguridad del punto de partida. Para hacer efectivamente seguro este punto de partida sugirió el método de la duda: aceptar sólo lo absolutamente indubitable.

Partió, pues, de su propia experiencia, que consideraba indubitable, ya que incluso dudar de ella misma parece presuponer la existencia del que duda (un sujeto que duda).

Yo no soy más escéptico que el propio Descartes acerca de la existencia de mí mismo, pero también pienso, como Descartes, que moriré pronto, lo cual no afectará mucho al mundo exceptuándome a mí mismo y a dos o tres amigos. Obviamente, tanto la propia vida como la muerte tienen algún significado, pero supongo (y creo que Descartes estaría de acuerdo) que mi propia existencia terminará sin que el mundo se acabe también.

Este punto de vista es de sentido común y constituye el credo fundamental de lo que se puede llamar "realismo". (Discutiremos pronto el realismo más extensamente.)

Admito que la creencia en la existencia propia es muy fuerte, mas no admito que pueda soportar el peso de algo semejante al edificio cartesiano; es excesivamente estrecha como plataforma de lanzamiento. Tampoco creo, dicho sea de paso, que sea tan indubitable como creía Descartes (lo cual es disculpable). En el maravilloso libro de Hugh Routledge, *Everest 1933*, leemos lo siguiente de Kipa, uno de los sherpas que subió a más altura de lo conveniente: "La pobre mente aturdida del pobre viejo Kipa aún se aferraba obstinadamente a la idea de que estaba muerto" [3]. No digo que la idea del pobre viejo Kipa fuese de sentido común ni siquiera razonable, pero, con todo, pone en duda la claridad e indubitabilidad que proclamaba Descartes. No propongo otra pretensión semejante de certeza, aunque admito de buena gana que es una prueba de sentido común bueno y sano el creer en la existencia del propio yo pensante. No trato de poner en tela de juicio la verdad del punto de partida de Descartes, sino su autosuficiencia en la tarea que se impuso, e incidentalmente, su indubitabilidad.

Locke, Berkeley e incluso el "escéptico" Hume y sus múltiples sucesores, especialmente Russell y Moore [4], compartían con Descartes la opinión de que las experiencias subjetivas son particularmente seguras y, por tanto, adecuadas como punto de partida o fundamento estable. Sin embargo, confiaban principalmente en experiencias de tipo observacional. También Reid, con quien comparto la adhesión al realismo y al sentido común, pensaba que tenemos una percepción de la realidad externa y objetiva muy directa, inmediata y segura.

Sugiero, por el contrario, que no hay nada directo o inmediato en nuestra experiencia: hemos de *aprender* que tenemos un yo que se prolonga en el tiempo y continúa existiendo incluso durante el sueño o la inconsciencia total y también hemos de aprender cosas sobre nuestro cuerpo y el de los demás. Se trata de descifrar o interpretar. Aprendemos tan bien a descifrar que todo se nos vuelve muy "directo" o "inmediato". Es lo mismo que ocurre con quien conoce el Morse o, para poner un ejemplo muy familiar, con quien sabe leer un libro: el libro le habla "directa", "inmeditamente". Sin embargo, sabemos que está funcionando un complicado proceso de descodificación. El aparente carácter directo e inmediato es el resultado del entrenamiento, como ocurre al tocar el piano o conducir un coche.

Hay razones para pensar que nuestras destrezas para descifrar tienen una base hereditaria. En cierta medida, cometemos algunas veces errores al

---

[3] Hugh Routledge, *Everest 1933*, Hodder & Soughton, Londres 1934, pág. 143. (Aunque sólo durante unos segundos, tuve una experiencia similar a la de Kipa una vez que me alcanzó un rayo en el *Sonnblick*, en los Alpes austríacos.)

[4] G. E. Moore era un realista porque amaba profundamente la verdad y veía con toda claridad la falsedad del idealismo. Desgraciadamente, creía en la teoría subjetivista del sentido común y, por ello, esperó en vano toda su vida encontrar una demostración del realismo basada en la percepción, tarea que resulta irrealizable. Russell, partiendo del realismo, incurrió de nuevo en el positivismo por la misma razón.

descifrar, sobre todo en período de aprendizaje y también más tarde, especialmente en situaciones poco comunes. El carácter inmediato y directo del proceso de descodificación bien aprendido no garantiza un funcionamiento sin faltas; no hay absoluta certeza, aunque sí una certeza suficiente para la mayoría de las tareas prácticas. Hay que abandonar la busca de la certeza, de una base segura para el conocimiento.

Así pues, veo el problema del conocimiento de un modo distinto que mis predecesores. No me preocupo del problema de la seguridad y justificación de las pretensiones de conocimiento, sino de su aumento. ¿En qué sentido podemos hablar de aumento o de progreso del conocimiento y cómo es posible realizarlo?

4. Realismo

El realismo es esencial al sentido común. El sentido común o el sentido común ilustrado distingue entre apariencia y realidad. (Se pueden poner ejemplos como "hoy el aire es tan diáfano que las montañas parecen mucho más cercanas de lo que realmente están" o, quizá, "parece que lo hace sin esfuerzo, pero me ha confesado que la tensión es casi insoportable".) Mas el sentido común también reconoce que las apariencias (la imagen reflejada en un espejo, por ejemplo) poseen una especie de realidad o, con otras palabras, que puede haber una realidad superficial —esto es, una apariencia— y una realidad profunda. Por otro lado, hay muchos tipos de realidades. El tipo más obvio es el de los alimentos (supongo que suministran las bases del sentimiento de realidad) o bien el de los objetos más resistentes (*objectum* = lo que se interpone en el curso de nuestra acción) como piedras, árboles o personas humanas. Pero hay otros tipos de realidad muy distintos como la descodificación subjetiva de nuestras experiencias sobre los alimentos, piedras, árboles y cuerpos humanos. El sabor y peso de los alimentos es también otro tipo de realidad, al igual que las propiedades de los árboles y los cuerpos humanos. Hay ejemplos de otros tipos en este universo tan variado como son, un dolor de muelas, una palabra, el lenguaje, un código de circulación, una novela o una decisión gubernamental; una demostración válida o inválida; tal vez, fuerzas, un campo de fuerzas, tendencias, estructuras y también regularidades. (Mis consideraciones dejan totalmente abierta la posibilidad de que estos diversos tipos de objetos se puedan relacionar entre sí, así como el modo en que lo hagan.)

5. Argumentos en favor del realismo

Sostengo que el realismo no es ni demostrable ni refutable. El realismo no es demostrable, al igual que todo lo que caiga fuera del campo de la lógica y la aritmética finita, pero mientras que las teorías científicas empí-

ricas son refutables [5], el relismo ni siquiera lo es. (Comparte esta irrefutabilidad con muchas teorías filosóficas o "metafísicas" y, en particular, con el idealismo.) Sin embargo, se puede argumentar, siendo abrumadores los argumentos en su favor.

El sentido común es un partidario acérrimo del realismo. Naturalmente, incluso antes de Descartes —de hecho, desde Heráclito— hay algunos asomos de duda acerca de si *nuestro mundo ordinario es algo más que un sueño.* Pero, hasta Descartes y Locke fueron realistas. Una teoría filosófica rival del realismo no surgió seriamente antes de Berkeley, Hume y Kant [6]. Kant, dicho sea de paso, incluso suministró una demostración del realismo, aunque no era válida. Considero importante, a este respecto, tener claro por qué no puede haber una prueba válida del realismo.

En su forma más simple, el idealismo afirma que el mundo (en el que se encuentra mi auditorio) no es más que un sueño. Ahora bien, está claro que esta teoría (aunque pensemos que es falsa) no es refutable: cualquier cosa que ustedes, mi auditorio, hagan para convencerme de su realidad —hablarme, escribirme una carta o incluso darme patadas— no puede tener la fuerza de una refutación, pues puedo seguir diciendo que sueño que ustedes me hablan, que he recibido una carta o una coz. (Se podría decir que todas estas respuestas son, de uno u otro modo, estratagemas inmunizadoras. Así es efectivamente, lo que constituye un poderoso argumento contra el idealismo; pero, una vez más, el que sea una teoría auto-inmunizadora no la refuta.)

Por tanto, el idealismo es irrefutable lo cual significa, obviamente, que es indemostrable. Sin embargo, estoy dispuesto a conceder que el realismo (como el idealismo) no sólo es indemostrable, sino también irrefutable; es decir, no se puede describir un suceso ni concebir una experiencia que constituyan una refutación efectiva del realismo [7]. En este caso, como en tantos

---

[5] Naturalmente, ésta es una de mis teorías más antiguas. Véase, por ejemplo, el capítulo 1 de mi libro *Conjectures and Refutations*, especialmente las págs. 37 y sigs. [Véase la versión castellana de Néstor Míguez, *El Desarrollo del conocimiento Científico. Conjeturas y Refutaciones*, Buenos Aires, Paidos, 1967; especialmente, las págs. 43 y sigs.] Estoy en desacuerdo con aquellos de mis críticos que afirman, por ejemplo. que la teoría de Newton no es más refutable que la de Freud. Tendríamos una refutación de la teoría newtoniana si, por ejemplo, todos los planetas continuasen moviéndose como ahora, mientras que la tierra se moviese en su órbita actual con aceleración constante, incluso al alejarse del perihelio. (Naturalmente, cualquier teoría se puede "inmunizar" —como dice Hans Albert— contra esta u otras refutaciones. Ya lo señalé en 1934, aunque no viene al caso repetirlo aquí.) He de decir que la refutabilidad de las teorías de Newton o Einstein es cuestión de física y metodología elementales. Einstein, por ejemplo, decía que su teoría de la relatividad general se vería refutada en caso de que no se observase el efecto de corrimiento hacia el rojo (el retraso de los relojes en campos gravitacionales fuertes) en las enanas blancas. Sin embargo, no hay descripción de una conducta humana lógicamente posible que resulte incompatible con las teorías psicoanalíticas de Freud, Adler o Jung.

[6] El positivismo, el fenomenalismo y la fenomenología están todos ellos, como es natural, impregnados del subjetivismo del punto de partida cartesiano.

[7] La irrefutabilidad del realismo, que estoy dispuesto a admitir, se puede poner en tela de juicio. La gran escritora austríaca Marie Ebner von Eschenbach (1830-1916) dice en unas memorias de su infancia que sospechaba que el realismo estaba equivocado. Tal vez las cosas desaparezcan cuando apartamos de ellas la mirada. Así pues,

otros, no habrá ningún argumento concluyente. *Con todo, hay argumentos en favor del realismo* o, más bien, *en contra del idealismo.*

(1) El argumento más fuerte puede consistir en una combinación de otros dos: (a) que el realismo forma parte del sentido común, y (b) que todos los pretendidos *argumentos* en su contra no sólo son filosóficos en el sentido más desacreditado del término, sino que además se basan en una parte del sentido común aceptada acríticamente, es decir, en esa parte errónea de la teoría del conocimiento del sentido común que ha denominado "la teoría de la mente como un cubo" (véase más adelante, secciones 12 y 13).

(2) Aunque hoy día la ciencia está un poco pasada de moda para algunas presonas, por razones que lamentablemente no son en absoluto nimias, no hemos de ignorar su relevancia para el realismo, aunque algunos científicos no sean realistas, como Ernst Mach o, en nuestros días, Eugene P. Wigner [8], cuyos argumentos caen de lleno en el tipo caracterizado en (1). Pero, olvidándonos por un momento de la física atómica (mecánica cuántica), podemos afirmar que casi todas —si no todas— las teorías físicas, químicas o biológicas implican el realismo en el sentido de que si son verdaderas, el realismo debe serlo también. Esta es una de las razones por las cuales algunas personas hablan de "realismo científico". Es una buena razón. Sin embargo, por mi parte, prefiero llamarlo "metafísico" antes que "científico" [9], dada su (aparente) falta de contrastabilidad.

---

intentaba sorprender la jugarreta del mundo en trance de desaparecer volviéndose rápidamnte con la vaga esperanza de ver cómo, a partir de la nada, las cosas se apresuraban a reunirse de nuevo. Con cada fracaso se veía a la vez defraudada y aliviada. Se pueden hacer muchos comentarios sobre esta historia. En primer lugar, es muy posible que este informe de experiencias infantiles no sea excepcional, sino normal y típico, desempeñando así un papel importante en el desarrollo de la distinción de sentido común entre apariencia y realidad. En segundo lugar (y me inclino ligeramente en su favor) es posible que el informe no sea típico porque la mayor parte de los niños, sean realistas ingenuos o se conviertan en tales a una edad demasiado temprana como para recordarlo. Ciertamente, Marie von Ebner no era un niño típico. En tercer lugar, no sólo en la infancia, sino también de adulto, he experimentado algo muy distinto de esto: por ejemplo, al encontrar algo que había olvidado totalmente, sentía a veces que si la naturaleza lo hubiese hecho desaparecer, nadie lo habría hecho con más habilidad. (No hacía falta que la realidad mostrase que existía "realmente", nadie se hubiese dado cuenta de que no lo hacía.) Si Marie hubiese tenido éxito, surge el problema de si eso habría refutado el realismo o si tan sólo habría refutado una forma muy especial del mismo. No me siento obligado a entrar en el problema, sino que por el contrario *concedo* a mis oponentes que el realismo es irrefutable. Si esta concesión fuese errónea, entonces el realismo estaría aún más próximo a una teoría científica contrastable de lo que yo pretendía en un principio.

[8] Para el caso de Wigner, véase especialmente su contribución a *The Scientist Speculates*, I. J. Good (ed.), Heinemann, Londres, 1962, págs. 284-302, y la crítica de Edward Nelson, *Dynamical Theories of Brownian Motion*, Princeton University Press 1967, 14-16. Véase también mi contribución en Mario Bunge (ed.), *Quantum Theory and Reality*, Springer, Berlín 1967, y en W. Yourgrau y A. van der Werde (eds.), *Perspectives in Quantum Theory, Essays in Honor of Alfred Landé*, M. I. T. Press 1971.

[9] Véase mi *logik der Forschung*, 1934 (*L. d F.*) donde, en la sección 79 (pág. 252 de la traducción inglesa, *The Logic of Scientific Discovery*, 1959—*L. Sc. D.* [Hay traducción castellana de Víctor Sánchez de Zavala, *La Lógica de la Investigación Científica*, Madrid, Tecnos, 1962, pág. 235]) me defino como un realista metafísico. En aquel entonces, identificaba incorrectamente los límites de la ciencia y los de la argumenta-

Mírese como se mire, hay excelentes razones para afirmar que *en la ciencia lo que se pretende es describir y (en la medida de lo posible) explicar la realidad*. Lo haremos mediante conjeturas teóricas; es decir, "teorías" que esperamos sean verdaderas (o próximas a la verdad), aunque no podamos demostrar su certeza ni siquiera su probabilidad (en el sentido del cálculo de probabilidades) a pesar de que sean las mejores teorías que podamos formular y de que, por tanto, puedan considerarse como "probables" en la medida en que este término está libre de toda asociación con el cálculo de probabilidades.

Hay otro sentido excelente y muy próximo en el que podemos hablar de "realismo científico"; el procedimiento adoptado entraña (en la medida en que no se vea destruido por actitudes irracionales) el éxito, en el sentido de que nuestras conjeturas teóricas tienden progresivamente hacia la verdad; es decir, hacia las descripciones verdaderas de ciertos hechos o aspectos de la realidad.

(3) Mas, aunque dejemos de lado los argumentos extraídos de la ciencia, quedan aún los argumentos lingüísticos. Toda discusión sobre el realismo, especialmente los argumentos en su contra, han de ser formulados en algún lenguaje. Ahora bien, el lenguaje humano es esencialmente descriptivo (y argumentador)[10] y una descripción sin ambigüedad siempre es realista: es *de* algo, de alguna situación que puede ser real o imaginaria. Así pues, si la situación es imaginaria, la descripción es sencillamente falsa y su negación constituye una descripción verdadera de la realidad en el sentido de Tarski. Esto no refuta lógicamente al idealismo o solipsismo, pero al menos lo hace irrelevante. Racionalidad, lenguaje, descripción, argumento, todos versan sobre alguna realidad y se dirigen a un auditorio. Todo esto presupone el realismo. Naturalmente, este argumento en favor del realismo no es lógicamente más concluyente que cualquier otro, pues puedo soñar sencillamente que estoy usando un lenguaje descriptivo y argumentos, mas, a pesar de todo, este argumento en pro del realismo es potente y *racional*. Es tan potente como la razón misma.

(4) Para mí, el idealismo es absurdo porque también implica que es mi mente la que crea este mundo tan hermoso, cuando sé de sobra que no soy su creador. Después de todo, la famosa frase "la belleza está en el ojo del espectador", aunque no sea explícitamente absurda, lo único que quiere decir es que la *apreciación* de la belleza es problemática. Me consta que la belleza

---

ción. Cambié de actitud más tarde, arguyendo que las teorías metafísicas no contrastables (i.e. irrefutables) pueden ser defendidas racionalmente. (Véase, por ejemplo, mi artículo "On the Status of Science and Metaphysics", publicado primero en 1958 y ahora en mi libro *Conjectures and Refutations*, 1963, 4.ª ed., 1972. [Cf. la traducción citada, capítulo 8, "sobre el Carácter de la Ciencia y la Metafísica", págs. 215 y siguientes.].)

[10] Bühler (anticipado en parte por W. von Humboldt) señaló claramente la función descriptiva del lenguaje. He aludido a ello en diversos lugares y he defendido la necesidad de introducir la función argumentadora del lenguaje. Véase, por ejemplo, mi artículo "Epistemología sin Sujeto Cognoscente" (leído en Amsterdam en 1967 y reimpreso ahora en este volumen en el capítulo 3).

de los autorretratos de Rembrandt no está en mis ojos, ni en mis oídos la de las Pasiones de Bach. Por el contrario, abriendo y cerrando mis ojos o mis oídos, puedo demostrar para mi satisfacción que no bastan para agotar toda la belleza que está ahí. También es cierto que otras personas son mejores jueces —más capaces que yo de apreciar la belleza de las pinturas y la música—. Negar el realismo equivale a la megalomanía (la enfermedad profesional más extendida entre los filósofos profesionales).

(5) De entre los muchos otros argumentos de peso (aunque inconcluyentes) que hay, sólo deseo señalar uno. Helo aquí: si el realismo es verdadero —especialmente, algo que se aproxime al realismo científico— entonces la razón de por qué es imposible probarlo es obvia. La razón estriba en que nuestro conocimiento subjetivo, incluso el perceptivo, consta de disposiciones a la acción y por eso es una especie de adaptación tentativa a la realidad. Además, a lo sumo, somos buscadores y, en todo caso, falibles. No hay garantía de no errar. Al mismo tiempo, todo el problema de la verdad y falsedad de nuestras opiniones y teorías pierde su sentido si no hay realidad, sino sólo sueños o ilusiones.

Resumiendo, propongo aceptar el realismo como la única hipótesis sensata —como conjetura a la que nunca se ha opuesto una alternativa sensata—. No quiero ser dogmático a este respecto, como a ningún otro, pero creo conocer todos los argumentos epsitemológicos —fundamentalmente, subjetivistas— que se han propuesto en favor de alternativas al realismo: el positivismo, el idealismo, el fenomenalismo, la fenomenología, etc. Aunque no me opongo a la discusión de los *ismos* filosóficos, considero que todos los *argumentos* filosóficos que han sido propuestos (que yo sepa) en favor de esa lista de *ismos* son, sin duda, erróneos. Algunos son resultado de la errónea búsqueda de la certeza o de fundamentos seguros sobre los que construir. Todos ellos son errores filosóficos típicos en el peor sentido de la palabra: todos derivan de una teoría del conocimiento equivocada, aunque de sentido común, que no resiste una crítica seria. (El sentido común naufraga irremisiblemente cuando se aplica a sí mismo; véase la sección 12, más adelante.)

Concluiré esta sección con la opinión de dos hombres que tengo por los más grandes de nuestro tiempo: Albert Einstein y Winston Churchill.

"No veo", escribe Einstein, "ningún peligro metafísico en aceptar las cosas —es decir, los objetos de la física... junto con las estructuras espacio-temporales que les corresponden" [11].

Esta era la opinión de Einstein tras un análisis simpatizante y cuidadoso de un brillante intento de refutar el realismo ingénuo debido a Bertrand Russell.

---

[11] Véase Albert Einstein, "Remarks on Bertrand Russell's Theory of Knowledge", en P. A. Schilpp (ed.), *The Philosophy of Bertrand Russell*, The Library of Living Philosophers, vol, V 1944, págs. 29 y sigs. La traducción de Schilpp de la página 291 es más literal que la mía, pero creo que la importancia de la idea de Einstein justifica mi intento de traducir *muy* libremente, pues creo que sigue siendo fiel a lo que Einstein quería decir.

Creo que las opiniones de Winston Churchill son reveladoras y constituyen un buen comentario sobre una filosofía que puede haber cambiado de color, pasando del idealismo al realismo, pero que sigue siendo tan inane como siempre: "Uno de mis primos que tuvo la gran suerte de recibir una educación universitaria", escribe, "solía darme la lata con argumentos para demostrar que nada existe si no pensamos en ello"; y prosigue:

> Siempre me tranquilizaba con el siguiente argumento que había ingeniado para mí mismo hace muchos años... ahí está el gran sol que, al parecer, no tiene más base que nuestros sentidos físicos. Pero afortunadamente, a parte de nuestros sentidos físicos, hay un método para comprobar la realidad del sol... los astrónomos... predicen con [las matemáticas] y la pura razón que una mancha negra lo atravesará un día determinado. Ustedes... miran, y el sentido de la vista indica inmediatamente que sus cálculos están justificados... *Tenemos lo que en cartografía militar se llama "un poste de referencia"*. Hemos logrado *un testimonio independiente* de la realidad del sol. *Cuando mis amigos metafísicos me dicen que los datos que emplean los astrónomos para sus cálculos se obtienen necesaria y primariamente mediante la evidencia de los sentidos, yo respondo, "No". Teóricamente, podrían obtenerse en cualquier caso con calculadoras automáticas puestas en marcha mediante la recepción de señales luminosas, sin intervención a ningún nivel de los sentidos humanos*... Yo... afirmo enfáticamente que el sol es real y que además está caliente, de hecho tan caliente como el infierno y, si los metafísicos lo dudan, que vayan y lo vean [1,2]".

Tal vez pueda añadir que pienso que el argumento de Churchill, en especial los importantes pasajes que he puesto en bastardilla, no sólo constituye una crítica válida a los argumentos idealistas y subjetivistas, sino que es además el argumento filosófico más sólido e ingenioso que conozco contra la epistemología subjetivista. No sé de ningún filósofo que haya prestado atención a este argumento (excepto algunos de mis estudiantes a los que llamé la atención sobre él). El argumento es tremendamente original; publicado por primera vez en 1930, es uno de los primeros argumentos filosóficos que hacen uso de la posibilidad de observatorios automáticos y máquinas calculadoras (programadas con la teoría newtoniana). Sin embargo, cuarenta años después de su publicación, Winston Churchill sigue siendo totalmente desconocido como epistemólogo: su nombre no aparece en ninguna de las muchas antologías sobre epistemología y ni siquiera aparece en la *Encyclopedia of Philosophy*.

Naturalmente, el argumento de Churchill se limita a ser una excelente refutación de los argumentos especiosos de los subjetivistas: *no demuestra el realismo*, porque los idealistas siempre pueden argüir que él o nosotros soñamos el debate con máquinas calculadoras y todo. Sin embargo, considero estúpido este argumento por su aplicabilidad universal. En cualquier

---

[1,2] Véase Winston S. Churchill, *My Early Life —A Roving Commission—*, publicado por primera vez en octubre de 1930. Citado con autorización de Hamlyn Publishing Group de Odhams Press ed., Londres 1947, capítulo IX, págs. 115 y sigs. (Las cursivas no son del original.) Véase también la edición de MacMillan, Londres, 1944, páginas 131 y sigs.

caso, a menos que algún filósofo construya un argumento totalmente nuevo, sugiero que en adelante ignoremos el subjetivismo.

6. Consideraciones sobre la verdad

Nuestra principal tarea filosófica y científica debe ser la búsqueda de la verdad. La justificación no es un objetivo, mientras que la brillantez y habilidad como tales son una pesadez. Deberíamos intentar ver o descubrir los problemas más urgentes, tratando de resolverlos proponiendo teorías verdaderas (o enunciados verdaderos o proposiciones verdaderas; no es necesario introducir distinciones) o, en cualquier caso, proponiendo teorías que se acerquen más a la verdad que las de nuestros predecesores.

Ahora bien, la búsqueda de la verdad sólo es posible si hablamos sencilla y claramente, evitando complicaciones y tecnicismos innecesarios. Para mí, buscar la sencillez y lucidez es un deber moral de todos los intelectuales: la falta de claridad es un pecado y la presunción un crimen. (La brevedad también es importante en vista de la explosión de las publicaciones, aunque es de menos importancia e incluso, a veces, incompatible con la claridad.) A menudo somos incapaces de llevar a cabo estas exigencias y no conseguimos decir las cosas clara e inteligiblemente, lo cual muestra simplemente que no somos suficientemente buenos como filósofos.

Acepto la teoría del sentido común (defendida y refinada por Alfred Tarski [13]) según la cual la verdad es la correspondencia con los hechos (o con la realidad) o, más exactamente, una teoría es verdadera si, y sólo si, corresponde a los hechos.

Para decirlo con algunos tecnicismos que gracias a Tarski se han hecho ya casi triviales: la verdad y la falsedad se consideran esencialmente como propiedades —o clases— de enunciados, es decir, de teorías o proposiciones ("oraciones significativas" [14]) formuladas sin ambigüedad en cierto lenguaje L_1 (por ejemplo, el alemán) sobre el que podemos hablar con toda libertad en otro lenguaje L_m, llamado también *metalenguaje*. Las expresiones de L_m que se refieren total o exclusivamente a L_1 se denominan "metalingüísticas".

Sea "*P*" uno de los nombres castellanos (L_m) de la expresión alemana (L_1) "*Der Mond ist aus grünem Käse gemach*". (Nótese que, mediante la

---

[13] Véase A. Tarski, *Logic, Semantics, Metamathematics*, Clarendon Press, Oxford 1956, págs. 152-278 (artículo publicado originalmente en polaco en 1933 y luego en alemán en 1936). También está publicado en *Philosophy and Phenomenological Research*, *4*, 1944, págs. 341-76. Véase más adelante el capítulo 9.

[14] La expresión "oración significativa" (i.e. una oración *más* su "significado"; esto es, un enunciado o proposición) pertenece a Tarski (en la traducción de Woodger). Tarski ha sido criticado injustamente por sostener la opinión de que la verdad es una propiedad de las (meras) oraciones: es decir, de las (asignificativas aunque) gramaticalmente correctas secuencias de palabras de un lenguaje o de un formalismo. La verdad es que Tarski, a través de su obra, sólo discute la verdad de los lenguajes *interpretados*. No distinguiré aquí entre enunciados, proposiciones, aserciones y teorías.

adición de comillas en castellano, la expresión alemana se ha convertido en un nombre metalingüístico castellano —llamado nombre citado— de la expresión alemana.) Así pues, la identidad "*P*='*Der Mond ist aus grünem Käse gemach*'" es evidentemente un enunciado metalingüístico castellano, con lo que podemos decir: "El enunciado alemán '*Der Mond ist aus günem Käse gemach*' corresponde a los hechos o al actual estado de cosas si, y sólo si, la luna está hecha de queso verde".

Introduzcamos ahora una regla general según la cual si *P* es un enunciado, entonces '*p*' es una abreviatura de la descripción castellana del estado de cosas al que se refiere el enunciado *P*. Entonces podemos decir, más en general: "la expresión *P* del lenguaje objeto es un enunciado que corresponde a los hechos si, y sólo si, *p*".

En castellano diríamos que "*P* es verdad en L_1" o "*P* es verdad en alemán". Sin embargo, la verdad no es un concepto relativo al lenguaje, porque si P_1 es un enunciado de un lenguaje, L_1, y P_2, de un lenguaje L_2, entonces vale lo siguiente (digamos, en L_m): si P_2 traduce a P_1 de L_1 a L_2, entonces P_1 y P_2 deben ser o ambos verdaderos o ambos falsos; han de tener el mismo valor de verdad. Si, además, el lenguaje es lo suficientemente rico como para poseer una operación de negación [15], entonces podemos decir que para todo enunciado falso, dicho lenguaje contiene otro verdadero. (De este modo, hablando grosso modo, sabemos que, en los lenguajes que disponen de una operación de negación, hay "tantos" enunciados verdaderos como falsos.)

La teoría de Tarski, en particular, deja claro *precisamente a qué hecho* corresponderá un enunciado *P*, si es que corresponde a algún hecho: a saber, al hecho de que *p*. También resuelve el problema de los enunciados falsos, ya que un enunciado falso, *P*, es falso *no* porque corresponde a cierta entidad extraña como un *no*-hecho, sino sencillamente porque *no* corresponde a *ningún* hecho: no mantiene con nada real la relación peculiar *corresponder* a un hecho, si bien mantiene una relación "descriptiva" con el espurio estado de hecho *p*. (No hay por qué evitar expresiones como "estado de cosas espurio" o incluso "hecho espurio", siempre que seamos conscientes de que sencillamente un hecho espurio no es real.)

Aunque hizo falta el genio de Tarski para ponerlo en claro, actualmente ya es perfectamente obvio que si queremos hablar sobre la correspondencia entre un enunciado y un hecho, precisamos un metalenguaje en que *enunciar* el hecho (o el hecho supuesto) al que se refiere el enunciado en cuestión, metalenguaje que además puede referirse a ese enunciado (mediante un *nombre* convencional o descriptivo de ese enunciado). Y viceversa:

---

[15] Aunque se hayan construido lenguajes artificiales sin la operación de la negación, que yo sepa, todos los lenguajes naturales disponen de ella. (Los que se dedican a la fisiología animal pretenden incluso que se puede encontrar algo parecido en las ratas que aprenden a apretar palancas con signos característicos y a comprender los símbolos que confieren a dichos signos un valor lógicamente opuesto al original. Véanse las referencias a R. W. Brown y K. L. Lashley en Hans Hörmann, *Psychologie der Sprache*, Springer, Berlín 1967, pág. 51.)

está claro que podemos hablar sobre la *correspondencia* entre enunciados y hechos siempre que dispongamos de un metalenguaje en el que poder hablar sobre (a) los hechos descritos por los enunciados de un lenguaje (objeto), por el simple expediente de enunciar estos hechos, y además (b) sobre los enunciados de este lenguaje (objeto), mediante la utilización de *nombres* de dichos enunciados.

Una vez enunciadas de este modo las condiciones en las que todo enunciado del lenguaje L_1 corresponde a los hechos, podemos establecer la siguiente definitición de un modo puramente verbal, aunque de acuerdo con el sentido común [16]: *Un enunciado es verdadero si, y sólo si, corresponde a los hechos.*

Como señala Tarski, esta noción de la verdad es objetiva o absolutista, aunque no sea absolutista en el sentido de permitirnos hablar con "absoluta certeza o seguridad", pues no nos suministra un criterio de verdad. Por el contrario, Tarski podía demostrar que, si L_1 es lo bastante rico (como para contener, por ejemplo, la aritmética), *entonces no puede haber un criterio general de verdad.* Sólo puede haberlo para lenguajes artificiales extremadamente pobres. En esto Tarski está en deuda con Gödel.)

Así, aunque la idea de verdad sea absolutista, no podemos pretender alcanzar una certeza absoluta: *somos buscadores de la verdad pero no sus poseedores* [17].

7. Contenido, contenido de verdad y contenido de falsedad

Para poner en claro lo que hacemos al buscar la verdad, hemos de poder dar razones, al menos en algunos casos, en favor de la pretensión intuitiva de que nos hemos *aproximado a la verdad* o de que una teoría T_1 ha sido superada por otra, digamos T_2, porque se parece más a la verdad que T_1.

Muchos filósofos, entre los que me encuentro, han recurrido intuitivamente a la idea de que una teoría, T_1, puede estar más alejada de la verdad que otra, T_2, de manera que ésta sea una mejor aproximación (o sencilla-

---

[16] Tarski muestra que, para evitar la paradoja del mentiroso, se precisa una precaución que va más allá del sentido común: hemos de tener cuidado de no emplear el término metalingüístico "verdadero (en L_1)" en el lenguaje L_1. Véase también más adelante el capítulo 9.

[17] El Profesor D. W. Hamlyn me ha hecho el gran honor de exponer mis opiniones en "The Nature of Science" (Paul Edwards —ed.— *The Encyclopedia of Philosophy*, vol. *3*, pág. 37). La mayor parte de su esbozo es totalmente correcta, pero me interpreta de un modo totalmente erróneo cuando resume mis puntos de vista diciendo que "la verdad misma no es más que una ilusión". Quienes no pretenden alcanzar una certeza absoluta sobre la paternidad de las comedias de Shakespeare o sobre la estructura del mundo, ¿se comprometen por ello con la doctrina de que el propio autor (o autora) de las comedias de Shakespeare o el propio mundo no es "más que una ilusión"?

(En diversos lugares de mis obras, especialmente en el capítulo 9 de este volumen, se podrá encontrar una explicación del gran significado que otorgo al concepto de verdad.)

mente mejor teoría) que T_1. Del mismo modo que muchos filósofos han considerado sospechosa la noción de verdad (no sin cierta dosis de verdad o razón, como ha dejado claro el análisis de las paradojas semánticas debido a Tarski), así también han desconfiado de la idea de mejor aproximación o acercamiento a la verdad, de mayor proximidad, o (como yo digo) de mayor "*verosimilitud*".

Para neutralizar estas sospechas he introducido una noción lógica de *verosimilitud*, combinando dos nociones introducidas originalmente por Tarski: (*a*) la noción de *verdad* y (*b*) la noción de *contenido* (*lógico*) de un enunciado; es decir, la clase de todos los enunciados implicados lógicamente por él (su "clase consecuente" [*consecuence class*], como acostumbra a llamarlos Tarski) [18].

Todo enunciado posee un contenido o clase consecuente, la clase de todos aquellos enunciados que se siguen de él. (Siguiendo a Tarski, podemos describir la clase consecuente de los enunciados tautológicos como la clase cero, de modo que los enunciados tautológicos tengan contenido cero.) Además, todo contenido posee un subcontenido que consta de todas sus consecuencias *verdaderas* y sólo de ellas.

La clase de los enunciados *verdaderos* que se derivan de un enunciado dado (o que pertenecen a un sistema deductivo dado) y que no son tautológicos se puede denominar su *contenido de verdad*.

El contenido de verdad de las tautologías (o enunciados lógicamente verdaderos) es cero: sólo consta de tautologías. Cualquier otro enunciado, *incluso los falsos*, no tiene un contenido de verdad cero.

La clase de los enunciados falsos implicados por un enunciado —la subclase de su contenido que consta exactamente de todos los enunciados falsos— se puede denominar (por cortesía, diríamos) su "contenido de falsedad", aunque no posee las propiedades características de un "contenido" o una clase consecuente de Tarski. No constituye un sistema deductivo de Tarski, ya que

---

[18] Aunque aquí no la discutiremos, es importante la diferencia entre el contenido o clase consecuente de un único enunciado o de un conjunto finito de enunciados (que siempre se puede sustituir por un único enunciado) por un lado, y una clase consecuente o contenido no axiomatizable (o no axiomatizable de modo finito) por otro. Las clases consecuentes de ambos tipos son denominadas por Tarski "sistemas deductivos": véase Tarski, *op. cit.* capítulo XII. Tarski introdujo varios años antes que yo la idea de clase consecuente. Llegué a esta idea, independientemente, más tarde en mi libro *Logik der Forschung*, en el que también introduje el concepto muy relacionado con éste de contenido empírico de un enunciado *S*, como clase de los enunciados empíricos incompatibles con *S* (o "prohibidos" por *S*). Carnap adoptó más tarde esta idea; véase especialmente su reconocimiento a mi *Logik der Forschung* en la página 406 de su libro *Logical Foundations of Probability*, 1950. La noción de *verosimilitud* la introduje en 1959 ó 60; véase la nota de la pág. 215 de *Conjectures and Refutations*, 3.ª ed., 1969 [dicha nota no aparece en la traducción castellana citada, hecha sobre la segunda edición, 1965]. He de señalar que mientras que en *Conjectures and Refutations* hablaba de "contenido-de-verdad" y "contenido-de-falsedad", ahora prefiero omitir los guiones cuando estos términos se usan como nombres (es decir, excepto en las raras expresiones tales como "medida contenido-de-verdad" [nosotros traducimos sencillamente "medida del contenido de verdad", utilizando únicamente la forma sustantiva]). En esto, sigo el consejo de Winston Churchill que aparece en la página 255 de la segunda edición del libro de Fowler, *Modern English Usage*, 1965.

de un enunciado falso es lógicamente posible deducir enunciados verdaderos. (La disyunción de un enunciado falso y cualquier otro verdadero constituye uno de esos enunciados verdaderos que se siguen del enunciado falso.)

En lo que queda de esta sección trataré de explicar las ideas intuitivas de contenido de verdad y falsedad un poco más detalladamente como preparación para una discusión más profunda de la idea de verosimilitud, ya que la *verosimilitud* de un enunciado se explica diciendo que *aumenta con su contenido de verdad y disminuye con su contenido de falsedad.* Para ello utilizaré ampliamente las ideas de Alfred Tarski, especialmente su *teoría de la verdad* y su *teoría de las clases consecuentes y de los sistemas deductivos* (aludidas ambas en la nota 18; para un tratamiento más detallado, véase también el capítulo 9 de este volumen).

Es posible explicar el contenido de falsedad de un enunciado *a* (en cuanto distinto de la clase de los enunciados falsos que se sigue de *a*) de tal modo que (*a*) sea un contenido (o una clase consecuente de Tarski), (*b*) contenga todos los enunciados falsos que se siguen de *a*, y (*c*) no contenga ningún enunciado verdadero. Para ello sólo precisamos relativizar el concepto de contenido, lo que se puede llevar a cabo de un modo muy natural.

Llamemos "*A*" al contenido o clase consecuente de un enunciado *a* (en general, *X* será el contenido del enunciado *x*). Llamenos con Tarski "*L*" al contenido de un enunciado lógicamente verdadero. *L* es la clase de todos los enunciados lógicamente verdaderos y el contenido común a todos los contenidos y a todos los enunciados. Podemos decir que *L* es el *contenido cero.*

Relativizamos ahora la idea de contenido de modo que podemos hablar del contenido relativo del enunciado *a*, dado el contenido *Y*, que denotamos mediante el símbolo "*a*, *Y*". Este es la clase de todos los enunciados deductibles de *a* en presencia de *Y* o con la ayuda de *Y*.

Se ve inmediatamente que si *A* es el contenido del enunciado *a*, entonces tenemos que $A = a, L$ (empleando la notación relativizada); es decir, el contenido absoluto *A* de un enunciado *a* es igual al contenido relativo de *a*, dada la "lógica" (= contenido cero).

Un caso más interesante del contenido relativo de una conjetura *a* es el caso a, B_t, donde B_t es nuestro *conocimiento básico* en el tiempo *t*; es decir, el conocimiento que se acepta sin discusión en un momento *t*. Podemos decir que lo interesante de una nueva conjetura *a* es, en primer lugar, el contenido relativo *a*, *B*; es decir, aquella parte del contenido de *a* que va más allá de *B*. Del mismo modo que el contenido de un enunciado lógicamente verdadero es cero, así el contenido relativo de una conjetura *a*, dado *B*, será cero si *a* sólo contiene conocimiento básico y nada más: en general, podemos decir que si *a* pertenece a *B* o, lo que es lo mismo, si $A \subset B$, entonces $a, B = 0$. Así, el contenido relativo de un enunciado *x*, *Y* es la información con que *x*, en presencia de *Y*, transciende *Y*.

Podemos definir ahora el contenido de falsedad de *a*, que simbolizamos mediante A_F, como el contenido de *a*, dado el *contenido de verdad* de *a*

(es decir, la intersección A_T de A y T, donde T es el sistema de Tarski de enunciados verdaderos). Es decir, podemos *definir*:

$$A_F = a,A_T.$$

El A_F así definido responde a nuestros deseos o condiciones de adecuación: (*a*) A_F es un contenido, aunque relativo; después de todo, también los contenidos absolutos son relativos, dada la verdad lógica (o suponiendo que L es lógicamente verdadero); (*b*) A_F contiene todos los enunciados *falsos* que se siguen de *a*, ya que es el sistema deductivo de enunciados que se siguen de *a*, tomando los enunciados *verdaderos* como *cero* (relativo); (*c*) A_F no "*contiene*" ningún enunciado verdadero en el sentido de que los enunciados verdaderos no se toman como contenido, sino como su contenido cero (relativo).

Los contenidos, unas veces son lógicamente comparables y otras no: forman sistemas parcialmente ordenados por la relación de inclusión del mismo modo que los enunciados forman sistemas parcialmente ordenados por la relación de implicación. Los contenidos *absolutos* A y B son comparables si suponemos que $A \subset B$ o $B \subset A$. Las condiciones de comparabilidad son más complicadas para los contenidos relativos.

Si X es un contenido axiomatizable de modo finito o un sistema deductivo, entonces existe un enunciado x tal que X es el contenido de x.

Así, si Y es axiomatizable de modo finito, podremos escribir

$$x,Y = x,y.$$

Ahora bien, en este caso vemos que x,Y es igual al contenido absoluto de la conjugación x. *y menos* el contenido absoluto de y.

Este tipo de consideraciones muestran que a,B y c,D serán comparables si

$$(A + B) - B \text{ es comparable a } (C + D) - D$$

donde "$+$" es la adición de Tarski de los *sistemas deductivos*: si ambos son axiomatizables, $A + B$ es el contenido de la conjunción *a.b*.

Por tanto, la comparabilidad no será frecuente en estos sistemas parcialmente ordenados. Pero hay un método para mostrar que estos sistemas parcialmente ordenados pueden ser ordenados linealmente "en principio" —es decir, sin contradición. El método consiste en la aplicación de la teoría formal de las probabilidades. (Aquí sólo mantengo la aplicabilidad para sistemas axiomatizables, aunque se puede ampliar a los no axiomatizables; véase, más adelante, el capítulo 9.)

Podemos escribir "$p(x,Y)$" o también

$$p(X,Y),$$

que se lee "la probabilidad de x dado Y", y aplicar al sistema axiomático formal la probabilidad relativa que he expuesto en otro lugar (por ejemplo,

en mi libro *Logic of Scientific Discovery*, nuevos apéndices *IV y *V) [19]. De ahí resulta que $p(x,Y)$ será un número entre 0 y 1 —normalmente no sabremos cuál —pudiendo afirmar con toda generalidad que

$$p(a,B) \text{ y } p(c,D) \text{ son } \textit{comparables en principio.}$$

Aunque normalmente no tengamos una información suficiente a nuestra disposición para decidir si

$$p(a,B) \leq p(c,D) \text{ o } p(a,B) \geq p(c,D),$$

podemos afirmar que al menos se mantiene una de las dos relaciones.

El resultado de todo ello es que podemos afirmar que los contenidos de verdad y falsedad se pueden hacer comparables en principio mediante el cálculo de probabilidades.

Como he mostrado en diversos lugares, el contenido A de a será tanto mayor cuanto menor sea la probabilidad lógica $P(a)$ o $p(A)$. La razón es que cuanta mayor información comporte un enunciado, menor será la probabilidad lógica de que sea verdadero (accidentalmente, como si dijéramos). Por tanto, podemos introducir una "*medida*" del contenido (se puede usar sobre todo topológicamente, es decir, como indicador del orden lineal),

$$ct(a),$$

es decir, el contenido (absoluto) de a, y también medidas relativas

$$ct(a,b) \text{ y } ct(a,B),$$

es decir, el contenido relativo de a dado b o B respectivamente. (Si B es axiomatizable, tenemos naturalmente $ct(a,b) = ct(a,B)$.) Estas medidas ct se pueden definir con ayuda del cálculo de probabilidades; es decir, con ayuda de la definición

$$ct(a,B) = 1 - p(a,B).$$

Tenemos ahora a nuestra disposición los medios para definir (la medida de) el contenido de verdad, $ct_T(a)$ y de falsedad $ct_F(a)$:

$$ct_T(a) = ct(A_T),$$

donde A_T es, una vez más, la intersección de A y el sistema de Tarski de todos los enunciados verdaderos; y

$$ct_F(a) = ct(a,A_T),$$

es decir, (la medida de) el contenido de falsedad es (la medida de) el contenido relativo de a, dado el contenido de verdad A_T de a; en otras palabras,

---

[19] Utilizé por primera vez una *medida* de contenido en 1954 (cf. *L. Sc. D.*, página 400 [trad. cit., pág. 372]) y medidas de contenido de verdad y falsedad, etc., en *C. & R.* (pág. 385 [trad. cit., pág. 443].). Tanto aquí como en el capítulo 9 señalo las *funciones de medida* con *minúsculas en cursiva* como *p, ct, vs.*

el grado en que *a* va *más allá* de los enunciados que (a) se siguen de *a* y (b) son verdaderos.

8. Consideraciones sobre la verosimilitud

Con ayuda de estas ideas podemos explicar ahora con mayor claridad lo que entendíamos intuitivamente por *verosimilitud*. Hablando intuitivamente, una teoría T_1 posee menos verosimilitud que una teoría T_2 si, y sólo si, (*a*) sus contenidos de verdad y falsedad (o sus medidas) son comparables y además (*b*) el contenido de verdad, pero no el de falsedad, de T_1 es menos que el de T_2 o también (*c*) el contenido de verdad de T_1 no es mayor que el de T_2, pero sí lo es el de falsedad. Resumiendo, diríamos que T_2 se aproxima más a la verdad o es más semejante a la verdad que T_1 si, y sólo si, se siguen de ella más enunciados verdaderos, pero no más enunciados falsos o, al menos, igual cantidad de enunciados verdaderos y menos enunciados falsos.

En general podemos decir que sólo teorías *rivales* (como las teorías sobre la gravitación de Einstein y Newton) son intuitivamente comparables respecto a sus contenidos (no medidos); pero también hay teorías rivales que no son comparables.

La comparabilidad intuitiva de los contenidos de las teorías newtoniana (*N*) y einsteiniana (*E*) se puede establecer del modo siguiente [20]: (*a*) para toda respuesta a un problema de la teoría newtoniana, hay una respuesta de la einsteiniana al menos de la misma precisión; esto hace que (la medida de) el contenido —en un sentido ligeramente más amplio que el de Tarski [21]— de *N* sea menor o igual que el de *E*; (*b*) hay problemas a los que la teoría de Einstein, *E*, puede suministrar una respuesta (no tautológica) que la teoría de Newton, *N*, no puede dar, lo que hace que el contenido de *N* sea claramente menor que el de *E*.

Así, podemos comparar intuitivamente los contenidos de ambas teorías,

---

[20] Este ejemplo lo discutí brevemente en la nota 7 de una contribución mía publicada por primera vez en *B. J. P. S., 5*, 1954, pág. 143 y sigs. y reimpresa en mi libro *Logic of Scientific Discovery*, segunda edición, 1968, nuevo apéndice IX; véase la pág. 401 [trad. cit., pág. 373]. Desde entonces, he reelaborado este punto. Véase, por ejemplo, mi artículo en honor a Herbert Feigl en P. Feyerabend y G. Maxwell (eds.), *Mind, Matter and Method*, 1966, págs. 343-53. En dicho artículo he mostrado que si los contenidos (no medidos) de dos teorías deductivas, *X* e *y*, son comparables, entonces sus contenidos de verdad son también comparables y son mayores o menores según lo sean los contenidos. Como ha mostrado David Miller, la demostración de este teorema se puede simplificar considerablemente. Es muy importante no olvidar lo siguiente: aunque las funciones *de medida* de los contenidos de verdad y de falsedad sean comparables *en principio* (ya que en principio las probabilidades son comparables), en general carecemos de medios para comparar, si no es comparando los contenidos no medidos de teorías rivales, precisamente de modo intuitivo.

[21] Originalmente, la noción de Tarski de una clase consecuente o contenido (no medido) nos permitía comparar los contenidos de las teorías cuando una de ellas implicaba la otra. La generalización presentada aquí nos permite comparar contenidos (o la medida de los contenidos) si una de ellas puede resolver todos los problemas que puede resolver la otra, al menos con la misma precisión.

teniendo la de Einstein mayor contenido. (Se puede mostrar que esta intuición se apoya en las medidas de contenido $ct(N)$ y $ct(E)$.) Esta situación hace que la teoría de Einstein sea *potencial* o *virtualmente* mejor, puesto que antes de cualquier contrastación podemos decir: si es verdadera, tendrá mayor poder explicativo. Además, nos desafía a emprender una mayor variedad de contrastaciones, con lo que nos ofrece nuevas oportunidades de aprender más sobre los hechos: sin el desafío de la teoría de Einstein nunca hubiéramos medido (con el elevado grado de precisión necesario) la distancia aparente entre las estrellas que están en las inmediaciones del sol durante un eclipse o el corrimiento hacia el rojo de la luz emitida por las enanas blancas.

Estas son algunas de las ventajas que posee, *incluso antes de ser contrastada,* una teoría (lógicamente) más potente (es decir, con mayor contenido) y que la convierten en una teoría potencialmente mejor o más desafiante.

Ahora bien, la teoría más potente, de mayor contenido, será también la de mayor verosimilitud *a menos que su contenido de falsedad sea también mayor.*

Esta afirmación constituye la base lógica del método de la ciencia —el método de conjeturas audaces seguidas de intentos de refutación. Una teoría será tanto más audaz cuanto mayor sea su contenido. También será tanto más arriesgada: para empezar, lo más probable es que sea falsa. Intentemos buscar sus puntos débiles, refutarla. Si no lo conseguimos o si las refutaciones que encontramos son a la vez refutaciones de la teoría más débil precedente [22], entonces tenemos razones para sospechar o conjeturar que la teoría más potente no tiene un contenido de falsedad superior al de su predecesora más débil y que, por tanto, tiene mayor grado de verosimilitud.

9. Verosimilitud y la búsqueda de la verdad

Representemos en un cuadrado la clase de todos los enunciados y dividámoslo en dos subáreas iguales para los enunciados verdaderos (T) y los falsos (F):

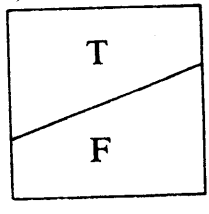


Fig. 1

[22] En cualquier caso, esto es lo que ocurre actualmente con el efecto del eclipse: las contrastaciones dan valores superiores a los predichos por E, mientras que N predice a lo sumo la mitad del resultado de E, incluso con la interpretación favorable de Einstein.

Cambiemos ahora la distribución, reuniendo la clase de los enunciados verdaderos en torno al centro del cuadrado.

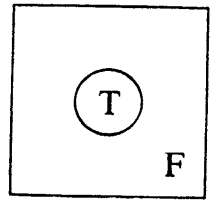


FIG. 2

La tarea de la ciencia es, metafóricamente hablando, acertar lo más posible en la diana (*T*) de los enunciados verdaderos (por el método de proponer teorías o conjeturas que parezcan prometedoras) y lo menos posible en el área falsa (*F*).

Es muy importante intentar hacer conjeturas que resulten teorías verdaderas, pero la verdad no es la única propiedad importante de nuestras conjeturas teóricas, puesto que no estamos especialmente interesados en proponer trivialidades o tautologías. "Todas las mesas son mesas" es ciertamente verdad —más ciertamente verdadero que las teorías de la gravitación de Einstein y Newton—, pero carece de interés intelectual: no es lo que andamos buscando en la ciencia. Wilhelm Busch compuso en una ocasión lo que he llamado una rima para la guardería epistemológica [23]:

Dos por dos son cuatro, es verdad,
pero también, demasiado vacío y simple.
Lo que busco es una clave
para cosas más difíciles.

En otras palabras, no sólo buscamos la verdad, vamos tras la verdad interesante e iluminadora, tras teorías que ofrezcan solución a *problemas* interesantes. Si es posible, vamos tras teorías profundas.

No nos limitamos a intentar hacer un blanco en un punto de la diana *T*, sino que procuramos cubrir un área lo más amplia e interesante posible: aunque sea verdad que dos por dos son cuatro, no constituye "una buena aproximación a la verdad" en el sentido aquí empleado, porque suministra demasiada poca verdad como para constituir, no ya el objeto de la ciencia,

---

[23] Del libro de Wilhelm Busch, *Schein und Sein*, 1909. El texto alemán es:
Zweimal zwei gleich vier ist Wahrheit,
Schade, dass sie leicht und leer ist.
Denn ich wollte lieber Klarheit
Über das, was voll und schwer ist.
Véase *Conjectures and Refutations*, pág. 230, nota 16 [trad. cit., pág. 267] y E. Nagel, P. Suppes y A. Tarski (eds.), *Logic, Methodology and Philosophy of Science*, Stanford U. P., 1962, pág. 290.

sino ni siquiera una parte suya importante. La Teoría de Newton es una "aproximación a la verdad" mucho mejor, aún cuando sea falsa (como probablemente sea), por la tremenda cantidad de consecuencias verdaderas interesantes e informativas que contiene: su *contenido de verdad* es muy grande.

Hay una cantidad infinita de enunciados verdaderos de muy distinto valor y, entre otros, un modo lógico de evaluarlos: estimamos el tamaño o medida de su *contenido* (que coincide con el contenido de verdad en el caso de los enunciados verdaderos, no en el de los falsos). El enunciado que suministre mayor información posee un contenido lógico o informativo mayor; es el mejor enunciado. Cuanto mayor es el contenido de un enunciado verdadero, mejor es como aproximación a nuestro blanco *T*, es decir, a la "verdad" (más exactamente, a la clase de todos los enunciados verdaderos), ya que no nos interesa aprender que todas las mesas son mesas. Cuando hablamos de aproximación o acercamiento a la verdad nos referimos a "toda la verdad", es decir, a toda la clase de enunciados verdaderos, la clase *T*.

Ahora bien, si un enunciado es falso, la situación es similar. Todo enunciado que no sea ambiguo es o verdadero o falso (aunque no sepamos cuál de las dos cosas es); la lógica que tengo aquí en cuenta [24] sólo posee estos dos valores, sin que exista una tercera posibilidad. Con todo, un enunciado falso puede parecer más próximo a la verdad que otro enunciado falso: "Ahora son las 9,45 p.m." parece más próximo a la verdad que "Ahora son las 9,40 p.m." si de hecho, cuando se hace la observación, son las 9,48 p.m.

No obstante, dicho así, la impresión intuitiva constituye un error: ambos enunciados son incompatibles y, por tanto, incomparables (a menos que introduzcamos una *medida* como *ct*). Sin embargo, hay algo de verdad en esta intuición errónea: si reemplazamos los dos enunciados por *enunciados de intervalo* (véase el párrafo siguiente), entonces el primero está más próximo a la verdad que el segundo.

Podemos proceder del siguiente modo; el primer enunciado se sustituye por "Ahora son entre las 9,45 p.m. y las 9,48 p.m." y el segundo por "Ahora son *entre* las 9,40 p.m. y 9.48 p.m.". De este modo, sustituimos cada enunciado por otro que admite un *rango* consecutivo *de valores*, un *rango de error*. Ahora ambos enunciados, así reemplazados, son comparables (ya que el primero implica el segundo) y además el primero está más próximo a la verdad que el segundo, lo que debe repecutir sobre cualquier función de medida de contenido que sea consistente, como *ct* y ct_T. Pero, puesto que en un sistema con una función de medida como ct_T se pueden comparar nuestros enunciados originales (en tal sistema todos los enunciados son comparables en principio), hemos de concluir que la medida del contenido de verdad ct_T se puede definir de manera que el ct_T del primer enunciado sea,

[24] Hay "sistemas polivalentes" de lógica con más de dos valores de verdad, pero son más débiles que los sistemas bivalentes, especialmente desde el punto de vista adoptado aquí (véase *Conjectures and Refutations*, pág. 64 [trad. cit., pág. 79]), según el cual la lógica formal es *el organon de la crítica*.

al menos, tan grande como —o mayor que— el del segundo enunciado, lo que hasta cierto punto justifica nuestra intuición original.

Nótese que la palabra "*entre*" que aparece en los enunciados de la sustitución se suele interpretar de modo que incluya o excluya los límites. Si lo interpretamos de manera que incluya los límites superiores, entonces ambos enunciados son verdaderos y en ambos casos $ct = ct_T$. Aunque son verdaderos, el primero tiene mayor verosimilitud porque tiene un mayor contenido de verdad que el segundo. Si, por otra parte, interpretamos "entre" de modo que excluya el límite superior, entonces ambos enunciados se tornan falsos (aunque pueda decirse que son "casi verdaderos") y continuamos afirmando que el primero se parece más a la verdad que el segundo. (Véase también mi *Conjectures and Refutations*, págs. 397 y sig., y mi *Logic of Scientific Discovery*, sec. 37. [Cf. la trad. cast., *Conjeturas y Refutaciones*, págs. 457 y sig. y *La Lógica de la Investigación Científica*, págs. 117 y sigs.].)

Así, sin violar la idea de la lógica bivalente ("todo enunciado que no sea ambiguo es o verdadero o falso"), podemos a veces hablar de enunciados falsos que son más o menos falsos o bien más alejados o más próximos a la verdad. Esta idea de mayor o menor verosimilitud es aplicable tanto a los enunciados falsos como a los verdaderos: la cuestión esencial es su *contenido de verdad*, concepto que cae de lleno en el campo de la lógica bivalente.

En otras palabras, parece que podemos identificar la idea intuitiva de *aproximación a la verdad* con la de *elevado contenido de verdad* y bajo "contenido de falsedad".

Esto es importante por dos razones: mitiga los recelos de algunos lógicos para operar con la idea intuitiva de aproximación a la verdad y nos permite decir que el objeto de la ciencia es la verdad, en el sentido de mejor aproximación a la verdad o mayor verosimilitud.

10. Verdad y verosimilitud como objetivos

Decir que el objeto de la ciencia es la verosimilitud, tiene considerables ventajas sobre la formulación, quizá más simple, de que el objeto de la ciencia es la verdad. Esto último puede sugerir que se alcanza totalmente el objetivo afirmando la indudable verdad de que todas las mesas son mesas o que $1 + 1 = 2$. Obviamente, ambos enunciados son verdaderos y tan obvio como esto es que ninguno de ellos se puede considerar un logro científico.

Además, los científicos buscan teorías como las de la gravedad de Newton o Einstein. Aunque estemos muy interesados en el problema de su verdad, éstas mantienen su interés aunque haya razones para creer que son falsas. Newton nunca creyó que su teoría fuese la última palabra ni Einstein que la suya fuese más que una buena aproximación a la teoría verdadera —la teoría del campo unificado que buscó desde 1916 hasta su muerte en 1955. Todo esto indica que la idea de "buscar la verdad" es satisfactoria sólo si (*a*) por "verdad" entendemos el conjunto de todas las

proposiciones verdaderas —es decir, nuestro inalcanzable conjunto que constituye la diana T (la clase de las proposiciones verdaderas de Tarski)— y (*b*) si aceptamos en nuestra investigación enunciados falsos como aproximaciones, con tal de que no sean "demasiado falsos" ("no tengan un contenido de falsedad demasiado grande") y tengan un gran contenido de verdad.

La búsqueda de la verosimilitud es, pues, una meta más clara y realista que la búsqueda de la verdad. Pero pretendo mostrar además que mientras que en las ciencias empíricas no podemos manejar argumentos suficientemente buenos como para pretender haber alcanzado efectivamente la verdad, con todo, podemos tener argumentos potentes y razonablemente buenos para pretender haber avanzado hacia la verdad; es decir, que la teoría T_2 es preferible a su predecesora T_1, al menos a la luz de todos los argumentos racionales conocidos.

Además podemos explicar el método de la ciencia, y gran parte de la historia de la ciencia, como el procedimiento racional de aproximarse a la verdad. (Se puede lograr otra clarificación importante con ayuda de la idea de verosimilitud en conexión con el problema de la inducción; véase especialmente la sección 32, más adelante.)

11. Comentarios en torno a las nociones de verdad y verosimilitud

Mi defensa de la legitimidad de la idea de verosimilitud ha sido a veces groseramente mal interpretada. Para evitar estas comprensiones defectuosas es recomendable tener presente que para mí no sólo son conjeturas las teorías, sino también las valoraciones de las teorías, incluso las comparaciones desde el punto de vista de su verosimilitud.

Es extraño que se haya malinterpretado este aspecto sumamente importante de mi teoría de la ciencia. Como he subrayado más de una vez, considero que toda valoración de teorías es *valoración del estado de su discusión crítica*. Por tanto, considero que la claridad es un valor intelectual, puesto que sin él la discusión crítica sería imposible. Pero no creo que la exactitud o precisión sean valores intelectuales en sí mismos; por el contrario, nunca trataremos de ser más exactos o precisos de lo que exige el problema que nos ocupa (que siempre consiste en discriminar entre teorías). Por esta razón, he subrayado que no me intereso por las definiciones; puesto que todas ellas han de emplear términos indefinidos, no importa, por regla general, que usemos un término como primitivo o como definido.

¿Por qué, pues, me he esforzado en mostrar que la verosimilitud se puede definir o reducir a otros términos (contenido de verdad, de falsedad y, en última instancia, probabilidad lógica)?

Algunas personas han supuesto que mi objetivo era fundamentalmente algo así como exactitud o precisión o, incluso, aplicabilidad: han supuesto

que deseaba encontrar una función numérica aplicable a teorías que nos dijese, en términos numéricos, cuál es su verosimilitud (o al menos su contenido de verdad o, tal vez, su grado de corroboración).

De hecho, nada hay más alejado de mis objetivos. Excepto en ciertos casos límites (como 0 y 1) no creo que se puedan determinar nunca numéricamente los grados de verosimilitud y las medidas del contenido de verdad o falsedad (o, digamos, del grado de corroboración e incluso de probabilidad lógica). Aún cuando la introducción de una función de medida haga comparables, en principio o en teoría, todos los contenidos, creo que a la hora de la aplicación efectiva tenemos que recurrir a esos pocos casos que son comparables, basándonos en consideraciones no-métricas y, como si dijéramos, cualitativas o lógicas en general, como en el caso de teorías *rivales* lógicamente más fuertes o más débiles; es decir, teorías que pretenden resolver los mismos problemas. Para comparar efectivamente tenemos que recurrir enteramente a estos casos (paradójicamente, podría decirse, porque las funciones de medida, como las probabilidades, hacen sus argumentos *generalmente* comparables en principio).

¿Qué pretenden, pues —se podría preguntar— mis intentos de mostrar que la verosimilitud es definibles en términos de probabilidad lógica? Mi objetivo es hacer con la verosimilitud (aunque con un grado inferior de precisión) lo mismo que hizo Tarski con la verdad: la rehabilitación de un concepto de sentido común que se ha hecho sospechoso, a pesar de que en mi opinión es absolutamente necesario para un realismo de sentido común crítico y para una teoría crítica de la ciencia. Es mi deseo poder decir que la ciencia tiene la verdad como fin, en el sentido de la correspondencia con los hechos o con la realidad. También es mi deseo decir (con Einstein y otros científicos) que la teoría de la relatividad es —o así lo suponemos— una mejor aproximación a la verdad que la teoría de Newton, del mismo modo que ésta constituye una mejor aproximación que la de Kepler. Además, es mi deseo poder decir estas cosas sin temor a que los conceptos de proximidad a la verdad o verosimilitud sean lógicamente incorrectos o "carentes de sentido". En otras palabras, pretendo rehabilitar una idea de sentido común que necesito para describir las metas de la ciencia y que subyace como principio regulador (aunque sólo sea de un modo inconsciente e intuitivo) a la racionalidad de toda discusión científica crítica.

Para mí, el mayor logro del descubrimiento que hizo Tarski de un método para definir la verdad (con respecto a lenguajes formalizados de orden finito) es la rehabilitación de la noción de verdad o correspondencia con la realidad, noción que se había hecho sospechosa. Al definirla en términos lógicos no sospechosos (no-semánticos), estableció su legitimidad. Una vez hecho ésto, mostró también la posibilidad de introducir por medio de axiomas una noción materialmente equivalente de verdad respecto a lenguajes formalizados de orden finito, si bien en este caso no se podía dar una definición explícita. En mi opinión, también rehabilitó el uso crítico

de la indefinida noción de verdad en lenguajes no formalizados ordinarios o de sentido común (que son de orden infinito), para lo cual basta con hacerlos ligeramente artificiales teniendo cuidado de evitar las antinomias. Diré que este lenguaje es de sentido común crítico: recuerdo que, en 1935, Tarski insistía con fuerza en la necesidad absoluta de emplear un lenguaje natural para construir un lenguaje formalizado, a pesar de que su uso acrítico nos conduzca a antinomias. Por tanto, tenemos que reformar, por así decir, el lenguaje ordinario a medida que lo usamos, tal como indicaba Neurath en su metáfora del barco que hay que reconstruir continuamente para tratar de mantenerlo a flote [25]. Esta es precisamente la situación del sentido común crítico tal como yo lo veo.

12. La equivocada teoría del conocimiento del sentido común

Sostengo que, aunque debe de ser criticado, el sentido común ha de ser siempre nuestro punto de partida. Por otra parte, como es de esperar, no vale gran cosa cuando reflexiona sobre sí mismo. De hecho, la teoría del sentido común acerca del conocimiento de sentido común no es más que un embrollo ingenuo. Sin embargo, ha suministrado los fundamentos sobre los que se han levantado incluso las más recientes teorías filosóficas sobre el conocimiento.

Esta teoría es muy simple. Si cualquiera de nosotros desea conocer algo que aún desconoce sobre el mundo, no tiene más que abrir los ojos y mirar a su alrededor. Hemos de enderezar las orejas y prestar oídos a los ruidos, especialmente los que hacen otras personas. Los diversos sentidos son, pues, nuestras *fuentes de conocimiento* —las fuentes o los accesos a nuestra mente.

Muchas veces me he referido a esta teoría llamándola la teoría de la mente como un cubo. La teoría del cubo se representa mejor mediante un diagrama:

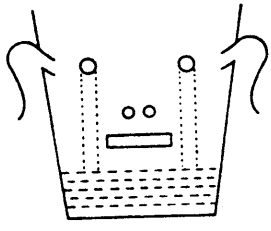


Fig. 3. El cubo

[25] Véase Otto Neurath *Erkenntnis, 3,* 1932, pág. 206. W. V. Quine nos ha recordado repetidas veces las consideraciones de Neurath; por ejemplo, en *Word and Object, M. I. T.* Press, 1960, pág. 3 [Traducción castellana de Manuel Sacristán Luzón, *Palabra y Objeto,* Barcelona, Labor, 1968, pág. 18.] o en *Ontological Relativity and Other Essays,* Columbia U. P., 1969, págs. 89 y 127.

Nuestra mente es en un principio un cubo, más o menos vacío, que se llena a través de los sentidos (o tal vez mediante un embudo para llenarlo desde arriba) y que almacena y digiere su contenido.

En el mundo filosófico esta teoría es conocida más dignamente con el nombre de teoría de la mente como *tabula rasa*: nuestra mente es *una pizarra vacía* en la que los sentidos gravan sus mensajes. Mas la característica principal de la teoría de la *tabula rasa* va más allá de la teoría de sentido común del cubo: me refiero al énfasis en el vacío completo de la mente en el momento de nacer. Para nuestros propósitos ésto no es más que una discrepancia secundaria entre ambas teorías, pues no importa el que nazcamos o no con algunas "ideas innatas" en nuestro cubo —más en el caso de los chicos inteligentes y menos en el de los retrasados—. La tesis importante de la teoría del cubo es que aprendemos la mayoría de las cosas, si no todas, mediante la entrada de la experiencia a través de las aberturas de nuestros sentidos, de modo que toda *experiencia consta de información recibida a través de los sentidos.*

Esta teoría totalmente errónea aún tiene mucha audiencia en la forma descrita. Aún sigue formando parte de teorías del aprendizaje o de la "teoría de la información", para poner algún ejemplo, si bien ahora se admite que el cubo no está vacío en el momento del nacimiento, sino dotado de un programa.

Mi tesis es que la teoría del cubo es manifiestamente ingénua y totalmente equivocados en todas sus versiones que, aceptadas inconscientemente de uno u otro modo, aún ejercen una influencia destructiva especialmente en los llamados conductistas, pues sugiere la aún poderosa teoría del reflejo condicionado y otras teorías que gozan de la más alta reputación.

He aquí algunos de los errores que contiene la teoría que concibe la mente como un cubo:

(1) Considera que el conocimiento consta de cosas o entidades de carácter cósico que se encuentra en nuestro cubo (tales como ideas, impresiones, sensaciones, datos de los sentidos, elementos, experiencias atómicas o —quizá mejor— experiencias moleculares o "*Gestalten*").

(2) Ante todo, el conocimiento está en nosotros: consta de informaciones que han llegado hasta nosotros y que hemos logrado absorver.

(3) Hay conocimiento *inmediato* o *directo;* es decir, los elementos puros e inadulterados de información que, habiendo llegado hasta nosotros, aún no están asimilados. Ningún conocimiento puede ser más elemental y cierto que éste.

El punto (3) se puede elaborar del modo siguiente:

(3ª) Según la teoría del sentido común, todo error, todo conocimiento equivocado proviene de una mala digestión intelectual que adultera estos elementos informativos últimos o "dados", malinterpretándolos o ligándolos equivocadamente entre sí. Las fuentes de error son nuestros propios

ingredientes subjetivos que añadimos a los elementos puros o dados de la información que, a su vez, no sólo están libres de error, sino que además son la norma de toda verdad hasta el punto de que estaría totalmente fuera de lugar plantear el problema de si son erróneos.

(3^b) Por tanto, el conocimiento, en la medida en que está libre de error, es un conocimiento que se recibe de un modo esencialmente pasivo, mientras que el error lo producimos siempre nosotros de manera activa (aunque no siempre de modo intencional), sea porque interferimos con lo dado, sea por cualquier otro tipo de inhabilidad: el cerebro perfecto nunca comete errores.

(3^c) El conocimiento que va más allá de la pura recepción de los elementos dados es, por tanto, menos cierto que el conocimiento elemental o dado que constituye el patrón de certeza. Si dudo de algo, no tengo más que abrir los ojos de nuevo y observar con mirada cándida, excluyendo todo prejuicio: he de purificar mi mente de las fuentes de error.

(4) Sin embargo, por razones de tipo práctico, precisamos conocer a un nivel más elevado, más allá de los simples datos o elementos. La razón es que necesitamos un conocimiento que establezca expectativas, conectando los datos existentes con los elementos inminentes. Este conocimiento superior se establece mediante *la asociación de ideas o elementos.*

(5) Las ideas o elementos se asocian cuando aparecen juntas y, lo que es más importante, *la asociación se refuerza con la repetición.*

(6) De este modo establecemos las *expectativas* (si la idea *a* está fuertemente asociada a la *b,* entonces la aparición de *a* hace esperar con gran fuerza la aparición de *b*).

(7) Las *creencias* surgen de manera similar. La creencia verdadera es la creencia en una asociación indefectible. La creencia errónea es la creencia en una asociación de ideas que, aunque tal vez hayan tenido lugar conjuntamente en el pasado, no aparecen siempre juntas indefectiblemente.

Resumiendo, lo que llamo teoría del conocimiento del sentido común se parece mucho al empirismo de Locke, Berkeley y Hume y no es muy ajeno al de muchos positivistas y empiristas modernos.

13. Crítica de la teoría del conocimiento del sentido común

Casi todo es erróneo en la teoría del conocimiento del sentido común, aunque tal vez el error central sea suponer que nuestra misión es lo que Dewey ha denominado *la busca de la certeza.*

Es esto lo que lleva a señalar los datos, elementos, e impresiones de los sentidos, o experiencias inmediatas, como base segura de todo conocimiento. Pero, lejos de ello, estos datos o elementos no existen en absoluto. No son más que invenciones de filósofos ilusos que han conseguido transmitirlas a los psicólogos.

¿Qué son los hechos? De niños aprendemos a descifrar los mensajes

caóticos que nos llegan del medio. Aprendemos a cribarlos, a ignorarlos en su mayor parte y a seleccionar aquellos que tienen para nosotros una importancia biológica inmediata o en el futuro, para el que nos vamos preparando mediante un proceso de maduración.

Aprender a descifrar los mensajes que nos llegan es algo extremadamente complicado que se basa en disposiciones innatas. Mi conjetura es que tenemos una disposición innata a referir los mensajes a un sistema coherente y en parte regular u ordenado: la "realidad". En otras palabras, nuestro conocimiento subjetivo de la realidad se compone de disposiciones innatas que van madurando. (En mi opinión, dicho sea de paso, esta construcción es demasiado sofisticada como para ser esgrimida como argumento independiente y potente en favor del realismo.) Sea lo que sea, aprendemos a descifrar por *ensayo y supresión de errores* y, aunque logremos una gran precisión y rapidez para experimentar el mensaje descifrado como si fuese "inmediato" o "dado", siempre aparecen algunos errores, corregidos normalmente mediante mecanismos especiales de gran complejidad y considerable eficiencia.

Por tanto, aunque forme parte del sentido común, toda esa historia de lo dado o de los datos verdaderos ligados a la certeza constituye una teoría errónea.

Admito que muchas cosas las experimentamos como inmediatamente dadas y como perfectamente ciertas, pero esto se debe a nuestro elaborado aparato descifrador con sus diversos ingenios construidos en el proceso de confrontación, que Winston Churchill llamaría "postes de referencia". Estos sistemas consiguen eliminar gran parte de los errores cometidos en el proceso de descifrado, de manera que rara vez nos equivocamos cuando experimentamos algo con carácter inmediato. Sin embargo, niego que estas experiencias bien adaptadas se puedan identificar en algún sentido con normas dadas de verdad o fiabilidad. De hecho, estos casos ni siquiera establecen una norma de "certeza" o de "claridad", ni demuestran que no podamos equivocarnos en nuestras percepciones inmediatas; ello se debe sencillamente a nuestra increíble eficiencia como sistemas biológicos. (Un fotógrafo experimentado rara vez hará exposiciones incorrectas debido a su entrenamiento y no a que sus fotografías hayan de tomarse como "datos", "normas de verdad" o tal vez como "normas de exposición correcta".)

Casi todos nosotros somos eficaces observando y percibiendo. Pero este problema hay que explicarlo recurriendo a teorías biológicas y no se puede tomar como base para ningún tipo de *dogmatismo* sobre el conocimiento directo, inmediato o intuitivo. Después de todo, todos fallamos alguna vez: nunca hemos de olvidar nuestra falibilidad.

14. Crítica de la teoría subjetivista del conocimiento

Naturalmente, todas estas cosas no refutan el idealismo ni la teoría subjetivista del conocimiento, ya que todo lo que he dicho sobre la psicología (o fisiología) de la percepción puede no ser más que un sueño.

Sin embargo, hay un argumento muy bueno que aún no he usado contra las teorías subjetivistas e idealistas. Es como sigue.

La mayor parte de los subjetivistas afirman con Berkeley que, a efectos prácticos, sus teorías concuerdan con el realismo y especialmente con las ciencias. Lo que ocurre, dicen, es que las ciencias no nos revelan las normas de verdad, no siendo más que instrumentos de predicción perfectos. No puede haber normas más elevadas de certeza (salvo la revelación divina) [26]. Pero entonces, la fisiología viene a decirnos que nuestros "datos" son falibles más bien que normas de verdad o certeza. Así, si esta forma de instrumentalismo subjetivo es verdadero, conduce a su propia refutación. Por tanto, no puede ser verdadero.

Naturalmente, esto no refuta al idealista que siempre podrá replicar que soñamos solamente que hemos refutado al idealismo.

Tal vez deba mencionar de pasada que es inaceptable un argumento, formalmente similar, de Russell contra el "realismo ingenuo" que impresionó mucho a Einstein. Helo aquí: "*Si hemos de creer* a la física (fisiología), el observador convencido de que está observando una piedra, *lo que observa en realidad son los efectos que la piedra produce en él.* De esta suerte, la ciencia parece estar en guerra consigo misma... El realismo ingenuo conduce a la física y ésta, si es verdadera, demuestra que el realismo ingenuo es falso. Por lo tanto, si el realismo ingenuo es verdadero, es falso; por tanto es falso" [27].

El argumento de Russell es inaceptable porque el pasaje que he puesto en cursiva constituye un error. *Cuando alguien observa una piedra, no observa los efectos que la piedra produce en él* (aunque tal vez sea así cuando contempla un dedo del pie herido), aún cuando descifra algunas de las señales que le llegan de la piedra. El argumento de Russell está al mismo nivel que este otro: "Cuando el lector cree estar leyendo a Russell, en realidad lo que observa son los efectos que Russell ejerce sobre él, por lo que no lee a Russell". Lo cierto es que leer (descifrar) a Russell se basa en parte en observaciones del texto de Russell; mas aquí no hay ningún problema digno de análisis, todos sabemos que leer es un proceso complejo en el que realizamos varios tipos de cosas a la vez.

No creo que valga la pena continuar estos ejercicios de agudeza, por

---

[26] Véase mi *Conjectures and Refutations*, capítulos 3 y 6.

[27] Cf. Bertrand Russell, *An Inquiry into Meaning and Truth*, Allen & Unwin, Londres 1940 (también New York), págs. 14 y sigs. (El subrayado es mío) [Traducción castellana de José Rovira Armengol, *Investigación sobre el significado y la verdad*, Buenos Aires, Losada, 1946, pág. 16]. Véase también el ensayo de Einstein en P. A. Schilpp (ed.), *The Philosophy of Bertrand Russell*, 1944, págs. 282 y siguientes.

lo que repito que aceptaré ingenuamente el realismo hasta tanto no se me ofrezcan nuevos argumentos.

15. Carácter pre-darwinista de la teoría del conocimiento del sentido común

La teoría del conocimiento del sentido común está radicalmente equivocada en todos sus puntos. Tal vez se puedan clarificar sus errores fundamentales del modo siguiente.

(1) *Hay conocimiento en sentido subjetivo que consta de disposiciones y expectativas.*

(2) *Pero también hay conocimiento en sentido objetivo, conocimiento humano, que consta de expectativas formuladas lingüísticamente sometidas a discusión crítica.*

(3) La teoría del sentido común no es capaz de ver que *la diferencia entre* (1) *y* (2) *tiene un gran alcance.* El conocimiento subjetivo no se puede criticar, aunque se pueda modificar por diversos métodos —por ejemplo, eliminando (matando) al portador de la disposición o conocimiento subjetivo en cuestión. Por tanto, el conocimiento en sentido subjetivo aumenta o se ajusta mejor mediante el método darwinista de mutaciones seguidas de supresión de organismos. Por el contrario, el conocimiento objetivo puede aumentar y cambiar mediante la eliminación (muerte) de las conjeturas lingüísticamente formuladas: el "portador" puede sobrevivir —incluso puede eliminar sus propias conjeturas, si es una persona autocrítica.

La diferencia estriba en que las teorías formuladas lingüísticamente se pueden *discutir críticamente.*

(4) Al margen de estos errores tan importantes, la teoría del sentido común está equivocada en diversos puntos. Es esencialmente una teoría sobre la génesis del conocimiento: la teoría del cubo versa sobre nuestra adquisición de conocimientos —en gran medida pasiva— con lo que también constituye una teoría de lo que he denominado el *aumento de conocimiento, aunque como teoría del aumento de conocimiento es manifiestamente falsa.*

(5) La teoría de la *tabula rasa* es pre-darwinista; toda persona que entienda algo de biología ha de tener claro el carácter innato de la mayoría de nuestras disposiciones, sea en el sentido de que hemos nacido con ellas (por ejemplo, la disposición a respirar, succionar, etc.) o en el sentido de que, en el proceso de maduración, el desarrollo de la disposición se ve solicitado por el medio (por ejemplo, la disposición a aprender un lenguaje).

(6) Pero aunque olvidemos toda teoría de la *tabula rasa* [28] y su-

[28] Se encontrarán algunos comentarios sobre la historia de la teoría de la *tabula rasa* en el nuevo apéndice sobre Parménides de la tercera edición de mi libro *Conjectures and Refutations,* 1969 y 1972 [La traducción castellana de la edición de 1965 no incluye dicho apéndice].

pongamos que el cubo está mediado al nacer o que cambia de estructura en el transcurso de la maduración, la teoría es aún peor. No sólo porque todo conocimiento subjetivo es disposicional, sino sobre todo porque no es una disposición de tipo asociativo (o del tipo del reflejo condicionado). Para decirlo clara y radicalmente: *no hay asociaciones ni reflejos condicionados.* Todos los reflejos son incondicionados; los supuestos reflejos "condicionados" son el resultado de modificaciones que eliminan total o parcialmente los comienzos falsos, es decir, los errores en el proceso de ensayo y error.

16. Bosquejo de una epistemología evolucionista

Que yo sepa, la expresión "epistemología evolucionista" se debe a mi amigo Donald T. Campbell. La idea es post-darwinista y se retrotrae hasta finales del siglo diecinueve, hasta pensadores como J. M. Baldwin, C. Lloyd Morgan y H. S. Jennings.

Aunque leí con gran interés durante los años precedentes a la redacción de mi primer libro, no sólo a Darwin, naturalmente, sino también a Lloyd Morgan y a Jennings, mi enfoque personal ha sido un tanto independiente de la mayoría de estas influencias. No obstante, como muchos otros filósofos, he hecho mucho hincapié en la distinción entre dos problemas del conocimiento: su génesis o historia, por un lado y los problemas de su verdad, validez y "justificación", por otro. (Así, por ejemplo, en el congreso de Praga de 1934 subrayé: "Las teorías científicas nunca se pueden 'justificar' o verificar, mas a pesar de ello, una hipótesis *A* puede dar mejores resultados en ciertas circunstancias que otra hipótesis *B*..." [**].) También subrayé hace mucho que los problemas de la verdad o validez, sin excluir *la justificación de la preferencia de una teoría a otra* (el único tipo de "justificación" que creo posible), ha de *distinguirse tajantemente de todo problema genético histórico y psicológico.*

Con todo, ya durante la redacción de mi libro *Logik der Forschung* llegué a la conclusión de que nosotros, los epistemólogos, podemos reclamar precedencia sobre los genetistas: las investigaciones lógicas sobre problemas de validez y aproximación a la verdad pueden ser de la mayor importancia para las investigaciones genéticas, históricas e incluso psicológicas. En cualquier caso, son lógicamente anteriores a este último tipo de problemas, aunque las investigaciones sobre historia del conocimiento pueden plantear importantes problemas al lógico de la investigación científica [**].

Hablo, pues, de epistemología evolucionista, aunque sostengo que las

---

[**] Véase *Erkenntnis, 5,* 1935, págs. 170 y sigs.; véase también mi *Logic of Scientific Discovery,* pág. 315 [trad. cit., pág. 294].

[**] A veces aludo al "principio de transferencia" para referirme al hecho de que lo que vale en lógica, vale en ciencia o en psicología, de manera que los resultados tienen aplicación psicológica o, en general, biológica. Véase la sección 4 de mi artículo "El conocimiento como Conjetura", capítulo 1 de este volumen, págs. 45 y siguientes.

ideas fundamentales en epistemología no son de carácter fáctico, sino lógico. A pesar de ello, todos sus ejemplos y la mayoría de sus problemas pueden ser sugeridos por estudios sobre la génesis del conocimiento.

Esta actitud es precisamente la opuesta a la de la teoría del sentido común y a la de la epistemología clásica, por ejemplo de Descartes, Locke, Berkeley, Hume y Reid: para Descartes y Berkeley la verdad queda garantizada por el origen de las ideas que, en última instancia, supervisa Dios. No sólo en Locke y Berkeley, sino también en Hume y Reid se encuentran trazas de la tesis de que la ignorancia es un pecado. Esto es así, porque el carácter claro o inmediato de nuestras ideas, impresiones y percepciones es el sello divino de su verdad que ofrece la mejor garantía al creyente. Mas, desde mi punto de vista, a veces consideramos que las teorías son verdaderas, incluso "inmediatamente" verdaderas, porque son verdaderas *y nuestro aparato mental está bien adaptado* a su nivel de dificultad. Sin embargo, nunca estamos "justificados" o "autorizados" a pretender que una teoría o creencia sea verdadera alegando su inmediatez o claridad. Para mí, esto es poner el carro delante de los bueyes: la inmediatez o claridad puede ser el resultado del hecho biológico de que una teoría sea verdadera, así como (en parte por esta razón) muy útil para nosotros. Sin embargo, argüir que la inmediatez o claridad establecen la verdad o son criterios de verdad es *el error fundamental del idealismo* [31].

Partiendo del realismo científico, está muy claro que no sobreviviremos si nuestras acciones y reacciones están mal ajustadas al medio. Puesto que las "creencias" están íntimamente ligadas a las expectativas y a la disposición a actuar, podemos decir que nuestras creencias más prácticas están próximas a la verdad en la medida en que sobrevivimos. Así se erigen en la parte más dogmática del sentido común que, aunque no sea en absoluto fiable, verdadero o cierto, constituye siempre un buen punto de partida.

No obstante, también sabemos que han desaparecido algunos de los animales más prósperos y que el éxito pasado no asegura en absoluto el éxito futuro. Esto es un hecho y, aunque siempre podemos hacer algo, nuestras posibilidades son muy limitadas. Menciono esto para dejar bien sentado que el éxito biológico pasado nunca asegura el éxito biológico futuro. Así, para un biólogo, el hecho de que una teoría haya tenido éxito en el pasado no garantiza en absoluto su éxito futuro.

---

[31] Para mí, el idealista epistemológico está en lo cierto al insistir en que todo conocimiento y el aumento del mismo —la génesis de la mutación de nuestras ideas— surgen de nosotros mismos y que sin estas ideas autogeneradas no habría conocimiento. El error está en no ver que sin la supresión de estas mutaciones mediante el enfrentamiento con el medio, no sólo no habría aliciente para ideas nuevas, sino que no habría ni siquiera conocimiento. (Cf. *Conjectures and Refutations*, especialmente, página 117 [trad. cit., pág. 139].) Kant estaba, pues, en lo cierto al decir que nuestro intelecto impone sus leyes —sus ideas, sus reglas— a la masa inarticulada de "sensaciones", poniendo orden en ellas. En lo que se equivocaba era en no ver que rara vez tenemos éxito con nuestra imposición, que ensayamos y erramos una y otra vez y que el resultado —nuestro conocimiento del mundo— se debe tanto a nuestras ideas autogeneradas como a la realidad que se resiste.

¿Cuál es la situación? Una teoría refutada en el pasado puede ser mantenida a pesar de su refutación. Así, podemos utilizar las leyes de Kepler para muchas cosas. Pero una teoría refutada en el pasado será *falsa*. No sólo buscamos éxito biológico o instrumental: en la ciencia *buscamos la verdad*.

He aquí un problema central de la teoría evolucionista: según esta teoría, los animales que no están bien adaptados a su medio ambiente perecen; en consecuencia, los que sobreviven (hasta un momento determinado) han de estar bien adaptados. Esta fórmula es casi tautológica, ya que "bien adaptado por el momento" equivale a "poseer las cualidades que le permiten sobrevivir hasta el momento". En otras palabras, una parte considerable del Darwinismo no posee el carácter de una teoría empírica, sino que es una *trivialidad lógica*.

Voy a aclarar lo que es y no es empírico en el Darwinismo. La existencia del medio con una cierta estructura es empírico. También lo es el que este medio cambie lentamente durante largos períodos de tiempo de un modo poco radical; si el cambio fuese demasiado radical, el sol podría explotar mañana dando lugar a una *nova*, con lo que desaparecería toda vida y toda adaptación de la superficie de la tierra. Dicho brevemente, la lógica no puede explicar la existencia en el mundo de condiciones en las que son posibles la vida y la lenta (sea lo que sea lo que signifique aquí "lenta") adaptación al medio.

Pero, dados unos organismos vivos sensibles a los cambios del medio y a las condiciones variables y suponiendo que no hay una armonía preestablecida entre las propiedades de los organismos y las del medio ambiente [32], podemos decir lo que sigue: los organismo sólo pueden sobrevivir si producen mutaciones, algunas de las cuales constituyen ajustes a los cambios que se avecinan, permitiendo así la mutabilidad. De este modo veremos que en la medida en que haya organismo vivos en un mundo cambiante, los que de hecho están vivos están perfectamente ajustados a su medio. Si el proceso de ajuste se ha producido durante mucho tiempo, el éxito, la sutileza y complejidad del ajuste puede sorprendernos por su aspecto milagroso. Sin embargo, puede decirse que el método de ensayo y supresión de errores que lleva a todo esto no es un método empírico, sino que pertenece a la *lógica de la situación*. Esto explica (tal vez demasiado brevemente) los elementos lógicos o *aprióricos* del darwinismo.

Ahora podemos ver de un modo más preciso que antes el tremendo avance biológico que representa la invención de un *lenguaje descriptivo y argumentador* [33]. La formulación lingüística de las teorías nos permite

---

[32] Tal vez sea interesante hacer aquí la siguiente consideración. K. Lorenz escribe en *Evolution and Modification of Behaviour*, Methuen, Londres 1966, págs. 103 y siguientes. [Traducción castellana de Carlos Gerhard, *Evolución y modificación de la conducta*, México, Siglo XXI, 1971, págs. 105 y sigs.]: "Toda modificabilidad que se revela regularmente como adaptativa, como lo hace indudablemente el aprendizaje, presupone una programación basada en información adquirida filogenéticamente. Para negar esto se necesita admitir una armonía preestablecida entre el organismo y el medio". Véase también la nota 34 más adelante.

[33] Para las diversas funciones del lenguaje humano, véase, por ejemplo, mi *Con-*

criticarlas y eliminarlas sin eliminar la estirpe que las sustenta. Esto constituye el primer logro. El segundo, es el desarrollo de una actitud crítica consciente y sistemática hacia nuestras teorías. Con esto comienza el método de la ciencia. La diferencia entre Einstein y una ameba, aunque ambos empleen el método de ensayo y supresión de errores, estriba en que a la ameba le desagrada equivocarse, mientras que a Einstein le intriga: busca errores conscientemente y desea aprender descubriéndolos y suprimiéndolos. El método de la ciencia es el método crítico.

Así la epistemología evolucionista nos permite comprender mejor tanto la evolución como la epistemología, en la medida en que coinciden con el método científico. Nos permite comprenderla mejor con una base lógica.

17. Conocimiento básico y problemas

El objeto de la ciencia es aumentar la verosimilitud. Como he mostrado, la teoría de la *tabula rasa* es absurda: en cada estadio de la evolución de la vida hemos de suponer la existencia de algún conocimiento bajo la forma de disposiciones y expectativas.

Según esto, *el aumento del conocimiento consiste en la modificación del conocimiento previo,* sea alterándolo, sea rechazándolo a gran escala. El conocimiento no parte nunca de cero, sino que siempre presupone un conocimiento básico —conocimiento que se da por supuesto en un momento determinado— junto con algunas dificultades, algunos problemas. Por regla general, éstos surgen del choque entre las expectativas inherentes a nuestro conocimiento básico y algunos descubrimientos nuevos, como observaciones o hipótesis sugeridas por ellos.

18. Todo conocimiento —incluso las observaciones— está impregnado de teoría

El conocimiento, en sus diversas formas subjetivas, tiene el carácter de disposiciones y expectativas. Consta de disposiciones de *organismos,* las cuales constituyen el aspecto más importante de su organización. Un tipo de organismo puede vivir hoy día sólo en el agua y otro sólo en tierra; puesto que han logrado sobrevivir hasta el presente, su misma ecología determina parte de su "conocimiento". Si no fuese absurdo hacer este tipo de estimaciones, diría que el 99,9 por 100 del conocimiento de un organismo es heredado o innato y sólo una décima parte consiste en modificaciones de dicho conocimiento innato. Sugiero, además, que también es innata *la plasticidad precisa* para estas modificaciones.

De aquí se sigue el teorema fundamental:

---

jectures and Refutations, págs. 134 y sigs. [trad. cit., págs. 158 y sigs.], y los capítulos 3, 4 y 6 más adelante.

Todo conocimiento adquirido, todo aprendizaje, consta de modificaciones (posiblemente de rechazos) de cierto tipo de conocimiento o disposición que ya se poseía previamente y, en última instancia, consta de disposiciones innatas [34].

De donde se sigue inmediatamente un segundo teorema:

Todo aumento de conocimiento consiste en el perfeccionamiento del conocimiento existente que se modifica con vistas a una mayor aproximación a la verdad.

Puesto que todas nuestras disposiciones son, en última instancia, ajustes a las condiciones del medio, invariables o de cambio lento, puede decirse que están *impregnadas de teoría,* entendiendo "teoría" en un sentido suficientemente amplio. Lo que quiero decir con esto es que toda observación está relacionada con un conjunto de situaciones típicas —regularidades— entre las que decidir. Creo que podemos afirmar aún más: *todos los órganos sensoriales incorporan genéticamente teorías anticipatorias.* El ojo de un gato reacciona de manera distinta a un número de situaciones típicas en vista de las cuales su estructura está dotada de mecanismos dispuestos y configurados en ella. Tales mecanismos corresponden a las situaciones biológicamente más importantes que ha de distinguir. Así pues, la tendencia a distinguir estas situaciones se configura en el órgano sensorial y con ella la *teoría según la cual estas situaciones, y sólo ellas, son las situaciones relevantes para cuya distinción ha de emplearse el ojo* [35].

El hecho de que todos nuestros sentidos estén de este modo impregnados de teoría muestra, de la manera más clara, el fallo radical de la teoría del cubo, así como de todas aquellas teorías que intentan remitir el conocimiento a las observaciones o al *input* del organismo. Por el contrario, *lo que se puede asimilar* (*y a lo que se puede reaccionar*) *como input relevante y lo que se ignora como irrelevante* depende completamente de la estructura innata (el "programa") del organismo.

19. Examen retrospectivo de la epistemología subjetivista

Desde el punto de vista aquí establecido, hemos de rechazar como totalmente infundada toda epistemología subjetivista que pretenda elegir como punto de partida eso que a ella no le parece en absoluto problemático; es decir, nuestras *experiencias observacionales "directas" o "inmediatas"*. Hay que admitir que, en general, estas experiencias son perfec-

[34] Para una defensa eficaz del conocimiento "innato" en contra de los conductistas y otros anti-teóricos, véase Konrad Lorenz, *Evolution and Modification of Behaviour,* citado en la nota 32.

[35] Véase, por ejemplo, los experimentos de T. N. Wiesel y D. H. Hubel, "Single-cell Responses in Striate Cortex of Kittens Deprived of Visión in One Eye". *Journal of Neurophysiology, 26,* págs. 1.003-17.

tamente "buenas" y eficaces (de lo contrario no habríamos sobrevivido), pero no son ni "directas" ni "inmediatas" ni mucho menos fiables.

No parece que haya ninguna razón para que no tomemos las experiencias observacionales como "punto de partida" provisional —punto de partida que, como el sentido común, no entraña un compromiso con la verdad o la certeza—. En la medida en que tengamos inclinaciones críticas, no importa mucho dónde o cómo empecemos. Sin embargo, partiendo de aquí (que tal vez sea lo que Russell llama "realismo ingenuo") llegamos, mediante la física y la biología, al resultado de que *nuestras observaciones son tremendamente complejas y no siempre fiables, a pesar de que constituyan procesos de descifrado asombrosamente excelentes* de las señales que nos llegan del medio. Por tanto, no deben de ser elevadas a la categoría de punto de partida como si fuesen una garantía de verdad.

Así pues, lo que se presentaba como una epistemología subjetivista o teoría de la *tabula rasa,* aparentemente libre de presupuestos, se desintegra totalmente. En su lugar, hemos de erigir una teoría del conocimiento en la que el sujeto cognoscente, el observador, desempeñe un papel importante aunque muy restringido.

20. Conocimiento en sentido objetivo

La teoría del conocimiento del sentido común y, con ella, todos los filósofos hasta Bolzano y Frege por lo menos, daban, equivocadamente, por supuesto que sólo había un tipo de conocimiento —el conocimiento poseído por algún sujeto cognoscente.

A este tipo de conocimiento lo denominaré "conocimiento subjetivo", a pesar del hecho de que, como veremos, *el conocimiento subjetivo genuino o inadulterado no existe.*

La teoría del conocimiento subjetivo es muy vieja, aunque se hace explícita con Descartes: "conocer" es una actividad que presupone la existencia de un sujeto cognoscente. *El que conoce es el yo subjetivo.*

Deseo distinguir ahora dos tipos de "conocimiento": el subjetivo (que habría que llamar mejor conocimiento organísmico, ya que consiste en disposiciones de organismos) y el objetivo o conocimiento en sentido objetivo que está formado por el contenido lógico de nuestras teorías, conjeturas, suposiciones (y, si lo deseamos, por el contenido lógico de nuestro código genético).

Son ejemplos de conocimiento objetivo las teorías publicadas en revistas y libros almacenados en bibliotecas, las discusiones de tales teorías, las dificultades o problemas señalados en relación con tales teorías [36], etc.

---

[36] He tratado estas cuestiones con mayor detalle en mis artículos "Epistemología sin Sujeto Cognoscente" (leído en Amsterdam en 1967) y "Sobre la teoría de la Mente Objetiva" (leído en Viena en 1968), reimpresos ambos en este volumen en los capítulos 5 y 6 respectivamente. Véase también la importante discusión de Sir John Eccles en su brillante libro *Facing Reality,* Springer, Berlín 1970, especialmente los

Podemos llamar al mundo físico "mundo 1", al mundo de nuestras experiencias conscientes, "mundo 2" y al mundo de los *contenidos* lógicos de los libros, bibliotecas, computadoras y similares, "mundo 3".

Tengo varias tesis acerca de este mundo 3:

(1) En el mundo 3 podemos descubrir nuevos problemas que estaban allí antes de ser descubiertos y antes de que se hiciesen conscientes; es decir, antes de que en el mundo 2 apareciese algo correspondiente a ellos. Ejemplo: descubrimos los números primos y de ahí se deriva el problema euclídeo de si la sucesión de los números primos es infinita.

(2) Por tanto, en algún sentido, el mundo 3 es *autónomo*: podemos hacer en este mundo descubrimientos teóricos del mismo modo que podemos hacer descubrimientos geográficos en el mundo 1.

(3) Tesis fundamental: casi todo nuestro conocimiento subjetivo (conocimiento del mundo 2) depende del mundo 3, es decir (al menos virtualmente), de las teorías *formuladas lingüísticamente*. *Ejemplos*: nuestra "auto-conciencia inmediata" nuestro "conocimiento de nosotros mismos", que es tan importante, depende en gran medida de las teorías del mundo 3: de nuestras teorías acerca de nuestro cuerpo y su existencia continua cuando quedamos dormidos o estamos inconscientes; de nuestras teorías acerca del tiempo (su carácter lineal); de nuestra teoría según la cual podemos evocar experiencias pasadas con diversos grados de claridad, etc. Con estas teorías se hallan conectadas nuestras expectativas de despertar tras haber quedado dormidos. Propongo la tesis de que *la plena conciencia de sí mismo* depende de todas estas teorías (del mundo 3) y de que los animales, aunque sean capaces de tener sentimientos, sensaciones, memoria y, por tanto, conciencia, no poseen la plena conciencia de sí mismos que constituye uno de los resultados del lenguaje humano y el desarrollo del mundo 3 específicamente humano.

21. La búsqueda de la certeza y la debilidad fundamental de la teoría del conocimiento del sentido común

La Teoría del conocimiento del sentido común pasa por alto el mundo 3, ignorando así la existencia del conocimiento en sentido objetivo. Esto constituye una gran debilidad de la teoría, aunque no es la más grave.

Para explicar cuál es la debilidad que me parece más grave, formularé primero dos enunciados, (a) y (b), característicos de esta teoría del conocimiento.

(a) El conocimiento es un tipo especial de creencia u opinión; es un estado especial de la mente.

(b) Para que un tipo de creencia o estado de la mente equivalga a

---

capítulos X y XI. David Miller ha llamado mi atención sobre la estrecha semejanza existente entre mi mundo 3 y el "tercer reino" de F. R. Leavis. Véase su conferencia *Two Cultures*, 1962, especialmente pág. 28.

algo más que a una "mera" creencia y pueda justificar la pretensión de ser un conocimiento, se precisa que el creyente esté en posesión de *razones suficientes* para establecer que el conocimiento en cuestión es *con certeza verdadero.*

De estas dos formulaciones, la (a) puede fácilmente ser reformulada de manera que pase a formar parte —una pequeña parte— de una teoría biológica del conocimiento que puede ser acetable. Podemos decir:

(a') El conocimiento subjetivo es un tipo de disposición de la que el organismo puede hacerse consciente algunas veces en forma de creencia, opinión o estado mental.

Este enunciado es perfectamente aceptable y podemos decir que se limita a exponer de un modo más exacto lo que (a) pretendía decir. Además, (a') es perfectamente compatible con una teoría del conocimiento que haga hincapié en el conocimiento objetivo; es decir, en el conocimiento como parte del mundo 3.

La situación de (b) es totalmente distinta. Tan pronto como tenemos en cuenta el conocimiento objetivo, hemos de decir que sólo podemos suministrar razones suficientes en favor de la verdad de una pequeñísima parte, a lo sumo, de dicho conocimiento: esta pequeña parte (si es que la hay) puede denominarse *conocimiento demostrable* y comprende (si es que comprende algo) las proposiciones de la lógica formal y la aritmética (finita).

Todo lo demás —con mucho, la parte más importante del conocimiento objetivo que comprende las ciencias naturales como la física y la fisiología— tiene esencialmente el carácter de conjeturas e hipótesis; sencillamente, no hay razones suficientes para sostener que estas hipótesis sean, no ya ciertamente verdaderas, sino ni siquiera verdaderas.

Así, (b) indica que si tratásemos de generalizar la teoría del conocimiento del sentido común para que abarcase el conocimiento objetivo, entonces sólo se podría admitir como tal el conocimiento demostrable (si lo hay). El vasto e importante campo de teorías que podemos llamar "conocimiento científico" no podría ser calificado de conocimiento en absoluto debido a su carácter de conjetura, ya que de acuerdo con la teoría del conocimiento del sentido común, el conocimiento es creencia *cualificada* —creencia cualificada de tal modo que es ciertamente verdadera. Es precisamente este tipo de cualificación lo que falta en el amplio e importante campo del conocimiento como conjetura.

Además, si el problema se enfoca desde el punto de vista de la teoría del conocimiento del sentido común, puede decirse que el término "conocimiento conjetural" es una contradicción en los términos. Pero la teoría del sentido común no es muy cabal en su subjetivismo. Por el contrario, no cabe duda de que en su origen la idea de "razón suficiente" era una idea objetivista: lo que originalmente se precisaba eran razones suficientes para probar o demostrar el conocimiento en cuestión, de modo que incluso (b) resulta ser una extensión de la idea objetivista de conocimiento

demostrable en el subjetivista mundo 2, el mundo de la disposición o "creencia". En consecuencia, toda generalización adecuada o traducción objetivista (b') habría de reducir, al igual que (a'), el conocimiento objetivo a conocimiento demostrable, con lo que habría que abandonar el conocimiento conjetural. Mas, con él, habría que abandonar también el conocimiento científico, el tipo de conocimiento más importante y el problema de toda teoría del conocimiento.

Creo que esto pone de manifiesto la debilidad fundamental de la teoría del conocimiento del sentido común. No sólo desconoce la distinción entre conocimiento objetivo y subjetivo, sino que además acepta, consciente o inconscientemente el conocimiento objetivo demostrable como paradigma de todo conocimiento, ya que sólo en ese caso poseemos realmente "razones suficientes" para distinguir el "conocimiento verdadero y cierto" de la "mera opinión" o la "mera creencia" [37].

No obstante, la teoría del conocimiento del sentido común sigue siendo esencialmente subjetivista. Por ello, incurre en la dificultad de admitir algo así como razones suficientes subjetivas, es decir, ciertos tipos de experiencia, creencia u opinión personales que, aunque subjetivas, son cierta e infaliblemente verdaderas, por lo que pueden pasar por conocimiento.

La dificultad es grande, puesto que, ¿cómo podemos discernir en el dominio de las creencias?, ¿cuáles son los criterios mediante los cuales reconocemos la verdad o la razón suficiente? O bien por la fuerza de la creencia (Hume), lo cual es difícil de defender racionalmente, o bien por su claridad y distinción, que se aducen (Descartes) como síntoma de su origen divino; o bien, más directamente, por su origen o génesis, es decir, por las "fuentes" de conocimiento. De este modo la teoría del sentido común se ve obligada a aceptar algún criterio de conocimiento "dado" (¿revelado?); se ve obligado a recurrir a lo dado por los sentidos o dato sensible, al sentimiento de inmediatez, intuibilidad o claridad. Es la pureza del origen lo que garantiza la liberación del error y, por tanto, la pureza del contenido [38].

Pero es evidente que todos estos criterios son espurios. El biólogo admitirá que nuestros órganos sensoriales son eficaces la mayoría de las veces y deberá explicar su eficacia mediante argumentos darwinistas. Con todo, negará que sean eficaces siempre o necesariamente y que podamos confiar en ellos como criterios de verdad. Su carácter "directo" e "inmediato" es

---

[37] Así pues, tenemos aquí el ejemplo de una estrategia típica de la teoría del conocimiento del sentido común: se toma una porción insuficiente de la lógica objetiva y se transfiere (tal vez inconscientemente) a la psicología. Es lo mismo que ocurre en la teoría de la asociación, en la que las dos "ideas" asociadas eran en su origen los "términos" de una proposición categórica cuya cópula es la asociación. (Piénsese en la "unión o separación de ideas" de Locke.)

[38] Para una descripción un tanto diferente de las fuentes del conocimiento y del problema del error, véase mi conferencia "On the Sources of Knowledge and of Ignorance", *Proceedings of the British Academy, 46,* 1960; véase también mi libro *Conjectures and Refutations,* 3.ª edición, 1969, págs. 3-30 ["Sobre las Fuentes del Conocimiento y de la Ignorancia" en trad. cit., págs. 9-40].

sólo aparente: no es más que otra faceta de la milagrosa facilidad y eficiencia con que trabajan, aún cuando en realidad trabajen de un modo muy indirecto mediante intrincados mecanismos de control construidos dentro del sistema.

Por tanto, no hay absoluta certeza en el terreno del conocimiento. Pero la doctrina (b) identifica la búsqueda de conocimiento con la búsqueda de la certeza, razón por la cual consituye la parte más débil de la teoría del conocimiento del sentido común.

Lo que hemos de hacer es partir del hecho de que el conocimiento científico objetivo tiene el carácter de una conjetura, para buscar después su análogo en el campo del conocimiento subjetivo. Este análogo se puede identificar fácilmente. Mi tesis es que el conocimiento subjetivo forma parte de un aparato de ajuste tremendamente complejo e intrincado, aunque asombrosamente preciso (en un organismo sano), que funciona esencialmente como el conocimiento conjetural objetivo: por el método de ensayo y eliminación de errores o mediante conjeturas, refutaciones y corrección propia ("autocorrección").

Al parecer, el sentido común forma parte de este aparato y, por ende, su situación *no es manifiestamente distinta* de la de otro conocimiento aparentemente "directo" o "inmediato". (En este punto Thomas Reid estaba en lo cierto, aunque encareció en gran medida la fuerza de los argumentos sacados de la inmediatez o claridad.)

22. Consideraciones analíticas sobre la certeza

No me interesan las definiciones ni el análisis lingüístico de las palabras o conceptos; sin embargo, se han dicho tantas cosas de tan escaso valor a cuento de la palabra "certeza" que hemos de decir aquí algo en aras de la claridad.

Hay una noción de certeza de sentido común que significa, brevemente, "suficientemente cierto a efectos prácticos". Cuando miro mi reloj, que es muy digno de confianza, y veo que señala las ocho en punto mientras oigo su tic-tac (señal de que no está parado), entonces tengo la "certeza razonable" o "certeza a efectos prácticos" de que son casi exactamente las ocho en punto. Cuando compro un libro y el librero me devuelve 20 peniques, tengo la "total certeza" de que las dos monedas no son falsas. (Mis "razones" son muy complejas: tienen algo que ver con la inflación que ha hecho que no les merezca la pena a los falsificadores acuñar monedas de diez peniques; a pesar de todo, las monedas en cuestión pueden ser piezas antiguas de los buenos tiempos en que merecía la pena falsificar florines.)

Si alguien me pregunta: "¿Está usted seguro de que la moneda que tiene en su mano es una moneda de diez peniques?", *tal vez la contemple de nuevo y diga*, "Sí". Pero si dependiesen muchas cosas importantes de la verdad de mi juicio, creo que me tomaría la molestia de ir al Banco más próximo para hacer que sometiesen la moneda a un examen detenido. Y,

si dependiese de ello la vida de un hombre, incluso intentaría llegar al cajero jefe del Banco de Inglaterra para pedirle que certificase el carácter genuino de la pieza.

¿Qué es lo que quiero decir con esto? Que la "certeza" de una creencia no es tanto un problema de intensidad cuanto de la *situación*: de nuestras expectativas acerca de sus consecuencias posibles. Todo depende de la importancia otorgada a la verdad o falsedad de la creencia.

Las "creencias" se relacionan con nuestra vida práctica diaria. *Actuamos basándonos en nuestras creencias.* (Un conductista diría: una "creencia" es algo en lo que nos basamos para actuar.) Por esta razón, en la mayoría de los casos basta un grado de certeza más bien bajo. Pero si dependen muchas cosas de nuestra creencia, entonces *no sólo* cambia *la intensidad* de la creencia, sino también su función biológica.

Hay una teoría subjetivista de la probabilidad que supone que podemos medir el grado de nuestra creencia en una proposición mediante las ofertas* que estaríamos dispuestos a aceptar al hacer una apuesta [39].

Esta teoría es increíblemente ingenua. Si deseo apostar y el dinero en juego no es mucho, puedo aceptar cualquier oferta. Si los riesgos son muy elevados, no acepto la apuesta de ninguna manera. Si no puedo dejar de apostar, digamos porque está en juego la vida de mi mejor amigo, puedo sentir la necesidad de asegurarme respecto a la proposición más trivial.

Con las manos en los bolsillos poseo la total "certeza" de que tengo cinco dedos en cada una; pero si la vida de mi mejor amigo dependiese de la verdad de esta proposición, podría sacar (y creo que lo haría) las manos de los bolsillos para estar "doblemente" seguro de que no he perdido milagrosamente alguno de los dedos.

¿Qué se sigue de todo esto? Se sigue que la "certeza absoluta" es una idea límite y que la "certeza" experimentada o subjetiva no depende solamente de los grados de creencia o de la evidencia, sino también de la situación —de la importancia de lo que está en juego. Además, incluso la evidencia en favor de una proposición que sé que es trivialmente verdadera puede ser radicalmente revisada, si lo que está en juego es suficientemente importante. Esto muestra que no es imposible mejorar incluso la más cierta de las certezas. La "certeza" no es una medida de la creencia— en sentido absoluto. Es más bien una medida de la creencia relativa a una situación inestable, ya que la importancia general de la situación en la que actuó tiene muchos aspectos y puedo pasar de unos a otros. Así, la certeza plena no posee el carácter de un máximo o un límite. Siempre puede haber una certeza aún más segura.

Dejando de lado las demostraciones *válidas y simples* en el mundo 3, la certeza objetiva sencillamente no existe. En el mundo 2 la certeza no

---

[39] Esta teoría se atribuye frecuentemente a F. P. Ramsey, aunque puede encontrarse en Kant.

* *Odds*: diferencia entre el dinero recibido en caso de ganar una apuesta y el pagado en caso de perderla. [*N. T.*].

es más que la sombra de una experiencia, la sombra de la fuerza de una creencia que depende no sólo de la "evidencia", sino de muchas otras cosas, tales como la seriedad de la situación problemática en la que actuamos (quizá no sea más que una cuestión de "temple").

En relación con esto es importante constatar que hay muchas situaciones en las que negarse a actuar equivale a una acción; en la vida diaria hemos de actuar constantemente y hemos de hacerlo basándonos en una certeza imperfecta (ya que difícilmente puede haber algo así como una certeza perfecta). Por regla general, la evidencia en base a la cual actuamos se acepta tras un examen de lo más superficial y *la discusión crítica de teorías rivales, característica de la buena ciencia, (en general) va más allá del tipo de cosas con las que nos conformamos en la vida práctica.*

(La ciencia —que es esencialmente crítica— tiene también un mayor carácter de conjetura y está menos segura de sí misma que la vida ordinaria, ya que hemos elevado conscientemente al rango de problema lo que normalmente formaría parte de nuestro conocimiento básico.)

Pero esto no quiere decir que alcancemos alguna vez una situación tal que un pensador científico ingenioso no pueda detectar brechas en nuestros argumentos: posibilidades que a nadie se le han ocurrido hasta el momento y que por tanto nadie ha intentado incluir o excluir.

Por tanto, desde el punto de vista del conocimiento objetivo, todas las teorías son conjeturas. Desde el punto de vista de la vida práctica, pueden ser mucho mejor discutidas, criticadas y contrastadas que todas las demás cosas que acostumbramos a tomar como base de acción y a considerar como ciertas.

No hay ninguna contradicción entre la tesis de que todo conocimiento objetivo es objetivamente conjetural y el hecho de que lo aceptamos en su mayor parte no sólo como "prácticamente cierto", sino como cierto en un sentido extraordinariamente cualificado; es decir, como mucho mejor contrastado que muchas teorías a las que confiamos constantemente nuestras vidas (tales como que el piso no se va a hundir o que no nos morderá una serpiente venenosa).

Las teorías son verdaderas o falsas y no *meros instrumentos*. Pero, naturalmente, también son instrumentos tanto para la práctica o la ciencia aplicada, como que para usted o para mí, cuando queremos decidir acerca de una teoría a la luz de su discusión crítica, que incluye informar acerca de sus contrastaciones. Si recibimos informes sobre los resultados de las contrastaciones y repetimos quizá nosotros mismos algunas de ellas, entonces podemos emplear estos informes y resultados para formar nuestras convicciones personales subjetivas y para determinar el grado de certeza con el que sostenemos nuestras creencias personales. (Este es uno de los modos en que se puede explicar el funcionamiento del principio de transferencia [40]: empleamos conocimiento objetivo en la formación de nuestras creencias personales subjetivas y, aunque las creencias personales subjetivas

[40] Cf. la sección 16, nota 30.

pueden considerarse siempre, en cierto sentido, como "irracionales", este uso del conocimiento objetivo muestra que no es necesario que se plantee aquí ningún conflicto humeano con la racionalidad.)

23. El método de la ciencia

He descrito tantas veces lo que considero el método autocorrector mediante el que avanza la ciencia, que puedo permitirme ser aquí muy breve: *El método de la ciencia es el método de conjeturas audaces e ingeniosas seguidas por intentos rigurosos de refutarlas.*

Una conjetura audaz es una teoría con gran contenido —mayor en cualquier caso que la teoría que, según deseamos, será superada por ella.

Que nuestras conjeturas hayan de ser audaces se sigue inmediatamente de lo que he dicho sobre el objetivo de la ciencia y la aproximación a la verdad: la audacia o gran contenido está ligado a un gran contenido de verdad, por lo que se puede ignorar, al principio, el contenido de falsedad.

Mas un aumento del contenido de verdad, en sí mismo, no basta para *garantizar* un aumento de verosimilitud. Puesto que el aumento de contenido es una cuestión puramente lógica y, puesto que el aumento de contenido de verdad viene dado junto con el aumento de contenido, el único campo que queda al debate científico —especialmente para las contrastaciones empíricas— consiste en ver si el contenido de fasedad ha aumentado también o no. Así, nuestra búsqueda competitiva de verosimilitud se convierte, especialmente desde el punto de vista empírico, en una comparación competitiva de los contenidos de falsedad (hecho que algunos consideran paradógico). Parece como si en la ciencia también valiese aquello de que las guerras nunca se ganan sino que siempre se pierden (como dijo en cierta ocasión Winston Churchill).

Nunca podemos tener la absoluta certeza de que nuestra teoría no esté perdida. Lo único que podemos hacer es buscar el contenido de falsedad de nuestra mejor teoría, cosa que llevamos a cabo intentando refutarla; es decir, intentando contrastarla de un modo riguroso a la luz de todos nuestros conocimientos objetivos y todo nuestro ingenio. Naturalmente, siempre cabe la posibilidad de que la teoría sea falsa aunque salga airosa de todas estas contrastaciones, lo cual viene dado por nuestra búsqueda de verosimilitud. *Pues si sale airosa de todas estas contrastaciones, podemos tener buenas razones para suponer que nuestra teoría, que como sabemos posee un contenido de verdad superior al de su predecesora, puede que no posea un contenido de falsedad mayor.* Además, si no logramos refutar la nueva teoría, especialmente en los dominios en que su predecesora ha sido refutada, entonces podemos tomar ésto como una de las razones objetivas en favor de *la conjetura de que la nueva teoría constituye una aproximación a la verdad mejor que la vieja.*

24. Discusión crítica, preferencia racional y el problema de la analiticidad de nuestras elecciones y predicciones

Vistas así las cosas, la contrastación de las teorías científicas forma parte de su discusión crítica o racional, como podemos decir, pues en este contexto, "crítico" es el mejor sinónimo que encuentro de "racional". La discusión crítica nunca puede producir razones suficientes para decir que una teoría es verdadera; nunca puede "justificar" nuestras pretensiones de conocimiento. Pero si tenemos suerte, la discusión crítica puede establecer razones suficientes para lo siguiente:

"En el momento presente, a la luz de una discusión crítica exhautiva y de contrastaciones rigurosas e ingeniosas, esta teoría puede ser con mucho la *mejor* (la más potente, la mejor contrastada); y por tanto, parece ser la que está más próxima a la verdad de todas las teorías rivales".

Para decirlo en pocas palabras: nunca podemos justificar racionalmente una teoría —es decir, la pretensión de que conocemos su verdad—, pero, si tenemos suerte, podemos justificar racionalmente la preferencia provisional de una teoría sobre todo un conjunto de teorías rivales; es decir, respecto al estado actual de la discusión. Aunque no podemos justificar la pretensión de que una teoría sea verdadera, podemos justificar el que a este nivel de la discusión todo indica que la teoría constituye *una aproximación a la verdad mejor* que cualquiera de las teorías rivales propuestas hasta este momento.

Consideremos dos hipótesis rivales h_1 y h_2. Mediante la abreviatura d_t nos referimos a la descripción del estado de la discusión de estas hipótesis en un momento t, incluyendo, naturalmente, la discusión de resultados experimentales y observacionales relevantes. Mediante.

$$c(h_1, d_t) < c(h_2, d_t) \tag{1}$$

denotaremos el enunciado que afirma que el *grado de corroboración* de h_1 es, a la luz de la discusión d_t inferior al de h_2. Indaguemos qué tipo de afirmación es (1).

De hecho, (1) será una afirmación un tanto incierta, al menos por la sencilla razón de que $c(h_1, d_t)$ cambia con el tiempo t, *pudiendo* cambiar con la rapidez del pensamiento. En muchos casos, la verdad o falsedad de (1) será sencillamente una cuestión de opinión.

Mas, supongamos circunstancias "ideales". Supongamos que una discusión prolongada ha llevado a resultados estables, especialmente a un acuerdo acerca de todos los elementos de juicio que hacen al caso y supongamos que no se producen cambios de opinión con t durante un período considerable.

En tales circunstancias, vemos que mientras que los elementos de juicio pertinentes de d_t son obviamente empíricos, el enunciado (1), suponiendo

que d_t sea suficientemente explícito, puede ser *lógico* o (a menos que nos disguste el término) "*analítico*". Esto estaría particularmente claro si $c(h_1, d_t)$ fuese negativo porque la discusión hubiese llegado, en el momento t, a un acuerdo acerca de que los elementos de juicio refutan h_1, mientras que $c(h_2, d_t)$ fuese positivo porque los elementos de juicio apoyan h_2. Ejemplo: sea h_1 la teoría de Kepler y h_2 la de Einstein. En el momento t podemos estar de acuerdo en que la teoría de Kepler está refutada (a causa de las perturbaciones newtonianas), mientras que los elementos de juicio que poseemos apoyan la teoría de Einstein. Si d_t es lo suficientemente explícito como para implicar todo esto, entonces

$$c(h_1, d_t) < c(h_2, d_t) \tag{1}$$

equivale a decir que un número negativo no especificado es menor que cierto número positivo también sin especificar, lo cual constituye un enunciado que puede ser denominado "lógico" o "analítico".

Naturalmente, podrían darse otros casos, como por ejemplo, si "d_t" no fuese más que un nombre como "el estado de la discusión el 12 de mayo de 1910". Pero del mismo modo que se puede decir que el resultado de comparar dos magnitudes conocidas es analítico, podemos decir también que el resultado de comparar dos grados de corroboración, si son conocidos de modo suficiente, es también analítico.

Mas sólo si el resultado de la comparación se conoce suficientemente bien, se puede decir que constituye la base de una preferencia racional; es decir, sólo si se sostiene (1) podemos decir que h_2 es racionalmente preferible a h_1.

Todo esto me parece patente y un tanto trivial, pero ha sido criticado por las razones siguientes.

Si (1) es analítico, entonces la decisión de preferir h_2 a h_1 también es analítica y, por tanto, *ninguna predicción sintética* nueva puede ir en contra de la preferencia de h_2 sobre h_1.

Aunque no estoy totalmente seguro, creo que lo que sigue resume las críticas, esgrimidas en primer lugar por el profesor Salmon, contra mi teoría de la corroboración: o todos los pasos descritos son analíticos —y entonces no puede haber predicciones científicas sintéticas— o hay predicciones científicas sintéticas y entonces algunos pasos no han de ser analíticos, sino genuinamente sintéticos o informativos y por tanto, inductivos.

Intentaré mostrar que este argumento no vale como crítica de mis opiniones: h_2 es sintético, como todo el mundo admite, y *todas las predicciones* (no tautológicas) se derivan de h_2 más bien que de la desigualdad (1). Basta esto para responder a la crítica. El problema de por qué preferimos h_2 a h_1 ha de resolverse —si está suficientemente especificado— con respecto a d_t, que tampoco es analítico.

Los motivos que han llevado a nuestra elección de h_2 no pueden alterar su carácter sintético. Estos motivos —en contra de los motivos psicológicos ordinarios— constituyen *preferencias racionalmente justificables.*

Por eso, en ellos desempeñan un papel la lógica y las proposiciones analíticas. Si lo deseamos, podemos decir que los motivos son "analíticos", mas dichos *motivos* analíticos para la elección de h_2 nunca hacen a h_2 *verdadera,* por no decir "analítica"; a lo sumo constituyen razones lógicamente inconcluyentes para *formular la conjetura* de que es la hipótesis más verosímil de todas las que compiten en el momento t.

25. Ciencia: el aumento del conocimiento mediante crítica e ingenio

Veo en la ciencia una de las mayores creaciones de la mente humana. Es una etapa comparable a la emergencia de un lenguaje descriptivo y argumentador o a la invención de la escritura. Es una etapa en la que los mitos explicativos se abren a la crítica consciente y consistente y en la que aceptamos el desafío de inventar nuevos mitos. (Es comparable a la etapa hipotética de los primeros momentos de la génesis de la vida en que los tipos de mutabilidad se hicieron objeto de evolución selectiva.)

Mucho antes de que apareciese la crítica había aumento del conocimiento —del conocimiento incorporado en el código genético—. El lenguaje permite la creación y mutación de mitos explicativos, lo que se ve potenciado considerablemente por el lenguaje escrito. Pero sólo la ciencia sustituye la eliminación del error, mediante la violenta lucha por la vida, por la crítica racional pacífica, lo que permite sustituir la muerte (mundo 1) y la intimidación (mundo 2) por los argumentos impersonales del mundo 3.

REFLEXION RETROSPECTIVA SOBRE LA INDUCCION

26. Los problemas de la causalidad y la inducción en Hume

Hasta ahora [41] he logrado hacer un bosquejo de la epistemología y los métodos empleados en la ciencia e incluso del aumento del conocimiento, sin mencionar siquiera la *inducción* —ni la palabra ni el supuesto fenómeno. Creo que es significativo. La inducción es un embrollo y, puesto

[41] Tal vez deba decir que el segundo capítulo, junto con esta "Reflexión Retrospectiva", lo escribí antes de las notas que han dado pie al capítulo 1 del presente volumen. Hay algunos solapamientos, como se deduce del hecho de que mi lema del capítulo 1 (sacado del libro de Bertrand Russell, *A History of Western Philosophy*, Londres, 1946, pág. 699) podría aparecer aquí perfectamente, en especial por lo que respecta a la sección 29. Sin embargo, los capítulos 1 y 2, sobre todo esta "Reflexión Retrospectiva", se complementan en varios aspectos.

que el problema de la inducción puede resolverse de un modo negativo aunque directo, resulta que la inducción no desempeña ningún papel epistemológico en la metodología de la ciencia o en el aumento del conocimiento.

En mi libro *Logik der Forschung* (1934) escribí: "Si siguiendo a Kant, llamamos problema de Hume al de la inducción deberíamos designar al problema de la demarcación como 'problema de Kant'"[42]. Que yo sepa, fue éste el primer pasaje en que se dio al problema de la inducción el nombre de "problema de Hume": el propio Kant no lo llamó así, en contra de lo que yo parecía dar a entender en el pasaje citado.

He aquí lo que ocurrió. Kant introdujo originalmente el nombre "problema de Hume" ("*Das Hume'sche Problem*")[43] para el problema de la condición epistemológica de la *causalidad* y, luego, lo generalizó al problema de si las proposiciones sintéticas podían ser válidas *a priori*, puesto que consideraba que el principio de causalidad era el más importante de todos los principios sintéticos válidos *a priori*.

Yo he procedido de modo distinto. Consideraba inútil el modo que tenía Hume de enfocar el problema de la causalidad. Se basaba en gran medida en su insostenible psicología empirista —su versión de la teoría de la mente como un cubo, cuyo contenido subjetivista y psicologista contenía pocas cosas que me pareciesen interesantes como contribución a la teoría del conocimiento objetivo. Pero, en medio de estas consideraciones subjetivistas, encontré una que me pareció una joya de valor incalculable para la teoría del conocimiento objetivo: una refutación lógica, sencilla y directa de toda pretensión según la cual la inducción constituye un argumento válido o un modo de razonar justificable.

Este argumento de Hume sobre la invalidez de la inducción constituía, al mismo tiempo, el meollo de su demostración de la no existencia de un nexo causal. Sin embargo, en cuanto tal, no lo encuentro ni muy relevante ni válido.

Así pues, a mi modo de ver, lo que Kant había denominado "problema de Hume", el problema de la causalidad, se divide en dos: el *problema causal* (respecto al cual estoy en descuerdo tanto con Kant como con Hume) y el *problema de la inducción*, respecto al cual estoy plenamente de acuerdo con Hume, en lo que atañe al aspecto lógico. (Hay también un aspecto psicológico del problema de la inducción respecto al cual, naturalmente, estoy en desacuerdo con Hume.)

El paso que di a continuación consistió en examinar con más detenimiento la situación problemática de Kant, encontrándome con que lo decisivo no era el principio de causalidad en cuanto uno de sus principios sin-

---

[42] Véase mi "Ein Kriterium des empirischen Charakters theoretischer Systeme", *Erkenntnis, 3*, págs. 426 y sigs., así como mi *Logic of Scientific Discovery*, sección 4 (tercer párrafo), 2.ª edición inglesa, 1968, pág. 34; 3.ª ed. alemana, 19-69, pág. 9 [trad. cit., pág. 34].

[43] I. Kant, *Prolegomena*. 1.ª ed., pág. 14 y sigs. [Traducción castellana de Julián Besteiro, Madrid, Aguilar, 1954].

téticos *a priori*, según él pensaba, sino el modo que tenía de usarlo, pues lo empleaba como un *principio de inducción*.

Hume mostró que la inducción no era válida porque llevaba a una regresión infinita. Ahora bien, a la luz del análisis kantiano (y mi rechazo de los principios sintéticos válidos *a priori*) he llegado a la siguiente fórmula: *la inducción no es válida porque lleva o a una regresión infinita o al apriorismo.*

Inicié la argumentación de mi libro *L.d.F.* con esta fórmula que me indujo a bautizar el meollo lógico de la cuestión —el problema de la inducción— con el nombre de "problema de Hume", atribuyéndoselo a Kant que había denominado así al problema de la causalidad (y su generalización).

Pero creo que debo entrar en más detalles, aunque sea brevemente.

Creo que Hume era un hombre de sentido común. Como señala en su *Treatise*, es un realista de sentido común convencido. Lo que hace de él un "escéptico" respecto a la realidad, conduciéndolo a esa forma radical de idealismo —el "monismo neutral", como lo llaman Mach y Russell— es precisamente esa parte mala de su sentido común: su teoría del conocimiento del sentido común, su versión de la teoría de la mente como un cubo. Hume, tal vez en mayor medida que Locke y Berkeley, es el paradigma del filósofo que, partiendo de un sentido común muy realista, se convierte —a causa de su teoría del conocimiento de sentido común— a una filosofía idealista que tiene por racionalmente insoslayable, aunque dividía su mente en dos partes. Es precisamente la esquizofrenia entre el realismo de sentido común y la teoría del conocimiento del sentido común la que hace desembocar al empirismo sensualista en un idealismo absurdo que sólo un filósofo puede aceptar, aunque ello le resulte difícil a alguien tan razonable como Hume.

Hume expresa del modo más claro esta esquizofrenia en el famoso pasaje:

"Como la duda escéptica surge naturalmente (según el sentido común) de una reflexión profunda e intensa sobre estos asuntos, aumenta siempre cuando llevamos más lejos nuestras reflexiones, ya sea en contra o en pro de ella. Tan sólo la falta de cuidado y atención pueden proporcionarnos algún alivio. Por esta razón confío enteramente en ellos y considero garantizado, cualquiera que sea la opinión del lector en el momento presente, *que de aquí a una hora estará persuadido de que existe un mundo externo e interno...*" [44].

Sin embargo, Hume estaba plenamente convencido de haber demostrado que su teoría del conocimiento era la más profunda y verdadera filosóficamente hablando. Para mostrar que era así efectivamente citaré el siguiente pasaje del *Treatise* (elegido entre una inmensidad de otros semejan-

---

[44] Hume, *Treatise*, Libro I, parte IV, sección II; Selby-Bigge, pág. 218 (el subrayado es mío). [Traducción castellana de Vicente Viqueira, *Tratado de la Naturaleza Humana*, Madrid, Calpe, 1923, tomo I, pág. 342.]

tes) en el que argumenta en contra del "error" [45] consistente en creer en un mundo externo:

> De todo ésto puede inferirse que no se requiere otra facultad más que los sentidos para convencernos de la existencia externa de los cuerpos. Sin embargo, para hacer imposible esta inferencia necesitamos tan sólo tener en cuenta las tres siguientes consideraciones: *Primera.* Propiamente hablando, no percibimos nuestro cuerpo cuando miramos nuestros miembros y partes, sino ciertas impresiones que nos dan nuestros sentidos; de modo que atribuir una existencia real y corporal a estas impresiones o a sus objetos es un acto del espíritu tan difícil de explicar como el que examinamos ahora. *Segunda.* Los sonidos, sabores y colores, aunque considerados por el espíritu como cualidades independientes continuas, no parecen tener una existencia en la extensión, y, por consiguiente, no pueden aparecerse a los sentidos como situados externamente al cuerpo. La razón de por qué les atribuimos un lugar será considerada después. *Tercera.* Aún nuestra vista no nos informa de la distancia o externidad, por decirlo así, inmediatamente y sin un cierto razonamiento y experiencia, como es reconocido por los filósofos más sensatos.

Es ésta la teoría del cubo en su estado puro: nuestro conocimiento consta de percepciones o "impresiones" que "*entran por los sentidos*" y que, una vez que constituyen conocimiento, han de estar en nosotros sin ningún tipo de separación o exterioridad.

(Naturalmente, esta profundidad filosófica no es más que un error. Una vez que arrancamos de la primera parte del sentido común, del realismo, vemos que somos animales provistos de órganos sensoriales que nos ayudan a descifrar las señales del mundo externo. Realizamos esta tarea asombrosamente bien con la cooperación prácticamente de todo nuestro cuerpo "exterior". Pero éste es otro problema del que no nos ocuparemos ahora.)

He bosquejado brevemente la esquizofrenia de Hume y el excesivo papel desempeñado en su filosofía por la teoría de la mente considerada como un cubo. En este contexto, voy a explicar ahora su teoría de la causalidad.

Esta teoría es compleja y dista mucho de ser consistente. Voy a subrayar solamente uno de sus aspectos.

Hume considera la causalidad (a) como *una relación entre eventos,* (b) como una "CONEXIÓN NECESARIA" (las mayúsculas son de Hume) [46].

Pero (dice) cuando "investigo el objeto de todos los sentidos para describir la naturaleza de esta conexión necesaria" no encuentro ninguna relación "más que... contigüidad y sucesión" [47]; no hay aquí ningún fundamento sensible para la idea de necesidad: tal idea carece totalmente de base.

Lo más próximo a esta idea, que es dado observar, es la *sucesión regular.* Pero si la sucesión regular de dos eventos fuese "necesaria", *tendría*

---

[45] Véase el *Treatise,* Libro I, parte IV, sección II, Selby-Bigge, págs. 190 y siguiente, el penúltimo párrafo [trad. cit., pág. 303].

[46] Hume, *Treatise,* Libro I, parte III, sección II; Selby-Bigge, pág. 77 [trad cit., página 133].

[47] Loc. cit.

que tener lugar con certeza no sólo entre casos observados, sino también entre casos inobservados. Es fundamentalmente de este modo cómo *el problema lógico de la inducción* hace aparición en la discusión subjetivista que hace Hume de la causalidad, en su búsqueda cubo-teórica del origen o el fundamento de la idea de necesidad.

Considero que este tipo de investigación está absolutamente mal planteada. Sin embargo, pienso que la formulación y tratamiento que hace Hume del *problema lógico de la inducción* (él nunca emplea esta expresión) es una joya intachable. Cito uno de los pasajes típicos:

"Haced que los hombres se persuadan una vez de estos dos principios: *que no existe nada en un objeto considerado en sí mismo que pueda proporcionarnos una razón para sacar una conclusión que vaya más allá de él, y que, aún después de la observación de la unión frecuente o constante de los objetos, no tenemos razón alguna para hacer una inferencia relativa a algún objeto remoto a estos de los que hemos tenido experiencia...* [48]." Estos "dos principios" de que Hume intenta persuadirnos contienen su *solución negativa del problema de la inducción.* Tales principios (y muchos otros pasajes similares) ya no hablan de causa o efecto, o de conexión necesaria. A mi juicio, constituyen las joyas lógicas enterradas en el barro psicológico del cubo. Precisamente para honrar a Hume por este descubrimiento fundamental cambié ligeramente el significado de la expresión kantiana "problema de Hume", ligándola al problema de la inducción más bien que al de la causalidad.

En este sentido, el problema lógico de la inducción de Hume es el problema de si podemos inferir casos inobservados a partir de casos observados, por muchos que sean; o enunciados "desconocidos" (no aceptados) a partir de enunciados "conocidos", por muchos que sean. La respuesta de Hume a este problema es claramente negativa; y, como señala, sigue siendo negativa aunque infiramos tan sólo la *probabilidad* de una conexión que no ha sido observada en lugar de su necesidad. Esta extensión a la probabilidad está formulada en el *Treatise*: "Según esta explicación de las cosas, que es, a mi parecer, de todo punto indiscutible, la probabilidad se funda en la presunción de una semejanza entre los objetos de los que tenemos experiencia y aquellos de los que no tenemos ninguna y, por consiguiente, es imposible que esta presunción surja de la probabilidad" [49].

Como se puede ver, el argumento en contra de la inducción probabilística es de carácter puramente formal. Esto se ve con mayor claridad aún en un pasaje del *Abstract* de Hume que he citado en mi libro *L. Sc. D.*, 1959 [50]. Es decir, Hume muestra que su argumento en contra de la vali-

---

[48] Hume, *Treatise*, Libro I, parte III, sección XII; Selby-Bigge, pág. 139 [trad. cit., pág. 228].

[49] Cf. *Treatise*, Libro I, parte III, sección VI; Selby-Bigge, pág. 90 [trad. cit., página 151].

[50] Cf. *L. Sc D.*, 1959 pág. 369 [trad. cit., pág. 344]: este pasaje se ocupa solamente de la inducción, mientras que el citado anteriormente del *Treatise*, pág. 91

dez de la inferencia inductiva se aplica lo mismo tanto si pretendemos inferir la "necesidad", *n*, de la conclusión, como si pretendemos inferir tan sólo su "probabilidad", *p*. (Las letras "*n*" y "*p*" serían variables que podrían intercambiarse en la argumentación de Hume [51].)

Además de este problema lógico de la inducción que Hume resolvió completamente (aún cuando su solución fuese negativa), hay otro problema lógico de la inducción que hay quien llama "problema de la inducción de Hume". Se trata del problema de cómo mostrar que las inferencias inductivas (al menos las probabilísticas) son, o pueden ser, válidas.

Este problema es típicamente pantanoso porque presupone acríticamente la existencia de una solución positiva de lo que llamo el "problema de Hume". Como el propio Hume ha mostrado, no existe una solución positiva.

Finalmente, tenemos el problema *psicológico* de la inducción de Hume. Puede formularse así: ¿Por qué la mayoría de las personas, por lo demás totalmente racionales, creen en la validez de la inducción? La respuesta de Hume es aquella a que alude Russell en el lema de nuestro capítulo primero: el mecanismo psicológico de asociación les fuerza a creer, por hábito o costumbre, que ocurrirá en el futuro lo que ha ocurrido en el pasado. Este mecanismo es útil biológicamente —tal vez no podríamos vivir sin él—, pero carece de todo fundamento racional. Así pues, no sólo es el hombre un animal irracional, sino que además es claramente irracional aquella parte de nosotros mismos que consideramos racional —el *conocimiento humano,* incluyendo el conocimiento práctico.

Así, la incompatibilidad de la solución negativa de Hume al problema de la inducción, con su solución positiva al problema psicológico, destruyó a la vez el empirismo y el racionalismo.

27. Por qué el problema lógico de la inducción de Hume es más profundo que su problema de la causalidad

Es fácil que se produzcan algunas pequeñas disputas sobre cuál de los dos problemas es más profundo: si el problema de la causalidad de Hume o lo que he denominado su problema de la inducción.

Se podría argumentar que si se resolviese positivamente el problema de la causalidad —si pudiéramos mostrar la existencia de un nexo necesario entre causa y efecto—, entonces el problema de la inducción se resol-

---

[trad. cit., pág. 152, al comienzo del párrafo en que está inserta la cita aludida], comienza con una discusión acerca de la causa y efecto.

[51] Hay un artículo de D. Stove, "Hume, Probability, and Induction", en *The Philosophical Review,* Abril 1965; reimpreso en *Philosophy Today, 3,* págs. 312-32, en el cual se discute mi tesis. Mas, Stove no puede estar en lo cierto, puesto que el argumento de Hume es de carácter formal (en el sentido de que arguye que nada cambia por el hecho de que sustituyamos *n* por *p*).

vería de manera positiva. Así pues, podría decirse que el problema de la causalidad es el más profundo.

Mi razonamiento sigue un camino inverso: el problema de la inducción se resuelve negativamente; nunca podemos justificar la verdad de la creencia en una regularidad. Ahora bien, empleamos regularidades constantemente como conjeturas, como hipótesis, y a veces poseemos buenas razones para preferir unas conjeturas a otras rivales.

En cualquier caso, a la luz de una conjetura podemos no sólo explicar mucho mejor que Hume la causa y el efecto, sino también decir en qué consiste el "nexo causal necesario".

Dada una conjetura acerca de una regularidad y unas condiciones iniciales que nos permitan derivar predicciones partiendo de nuestra conjetura, podemos llamar causa (conjeturada) a las condiciones y, efecto (conjeturado) al evento predicho. Finalmente, la conjetura que los une con necesidad lógica constituye el nexo necesario (conjeturado) entre causa y efecto tanto tiempo buscado. (El conjunto puede ser denominado "explicación causal" como ya hice en *L.d.F.*, sección 12. [Traducción citada, capítulo III, sección 12, págs. 57-60, especialmente pág. 58.])

Esto indica que la solución negativa dada por Hume al problema de la inducción nos lleva mucho más lejos que su solución negativa al problema de la causalidad, de manera que podemos decir que el primero es "más profundo" o que "subyace" al segundo.

28. Intervención kantiana: conocimiento objetivo

Kant se percató de que la solución negativa dada por Hume al problema de la inducción destruía la racionalidad de los fundamentos de la dinámica newtoniana. Kant, como todos sus contemporáneos cultos, no dudaba de la verdad de la teoría de Newton. Sin embargo, el análisis de Hume la reducía a "costumbre" o "hábito", posición totalmente inaceptable.

Hume había mostrado que la inducción se veía amenazada por una regresión infinita. Kant señaló que Hume, con su dogmatismo empirista, no había tenido en cuenta la posibilidad de que hubiese un principio de causalidad (mejor dicho, un principio de inducción) válido *a priori*. Esta fue la posición adoptada por Kant (como ya expliqué en la sección 1 de *L.d.F.* [*Ibid.*, págs. 27-30]) y que Bertrand Russell adoptó también posteriormente: ambos intentaron salvar la racionalidad humana del irracionalismo de Hume.

Kant dividía las oraciones según su forma lógica en analíticas y sintéticas, siendo las analíticas aquellas cuya verdad o falsedad puede determinarse con la sola ayuda de la lógica. A continuación, introdujo la división según su validez *a priori* o *a posteriori*: según que su pretendida verdad o falsedad no precisase un fundamento empírico *(a priori)* o sí lo precisase *(a posteriori)*.

Puesto que, por definición, todos los enunciados analíticos eran *a priori*, llegamos al siguiente cuadro:

División de los enunciados

		Según su forma lógica	
		analíticos	sintéticos
Según el fundamento basis of claim to truth or falsity:	*a priori*	+	?
	a posteriori →	−	+

(Las flechas significan "si... entonces"; por ejemplo: si analítico, entonces *a priori*.)

El cuadro señala que analítico implica *a priori* y *a posteriori* implica sintético. Mas esto deja abierto el problema de si hay o no enunciados sintéticos que puedan ser válidos *a priori*. Kant afirmaba que sí, pretendiendo que la aritmética, la geometría, el principio de causalidad (y gran parte de la física newtoniana) eran sintéticos y válidos *a priori*.

Para Kant esto resolvería el problema de Hume. ¿Pero, acaso constituye una teoría sostenible? ¿Cómo se puede establecer *a priori* la verdad del principio de causalidad (por ejemplo)?

Es aquí donde Kant emprendió su "Revolución copernicana": *el entendimiento humano inventa e impone sus leyes a la ciénaga sensual, creando de ese modo el orden de la naturaleza.*

Era ésta una teoría audaz. Sin embargo, se vino abajo al descubrirse que la dinámica newtoniana no era válida *a priori*, sino que era una hipótesis maravillosa —una conjetura.

Desde el punto de vista del realismo de sentido común se podía salvar una pequeña parte de la idea kantiniana. Las leyes de la naturaleza *son* invención nuestra, son genéticamente *a priori* por construcción animal o humana, aunque no son válidas *a priori*. *Intentamos* imponerlas a la naturaleza y con mucha frecuencia fracasamos, pereciendo junto con nuestras conjeturas equivocadas. Sin embargo, otras veces nos aproximamos a la verdad lo suficiente como para sobrevivir con nuestras conjeturas. Además, a nivel humano, podemos criticar sistemáticamente nuestras conjeturas tan pronto como tenemos a nuestra disposición un lenguaje descriptivo y argumentativo. Este es el método de la ciencia.

Es importante constatar la gran contribución kantiana a esta solución, a pesar de que él mismo no superó completamente el subjetivismo en la teoría del conocimiento. Tal vez el paso más importante fuese su empeño en discutir constantemente los enunciados, proposiciones, principios y teo-

rías científicas junto con los argumentos en pro y en contra, allí donde sus predecesores habían hablado fundamentalmente de sensaciones, impresiones o creencias.

29. La solución de la paradoja de Hume: restauración de la racionalidad [52]

Después de haber redactado el pasaje en que he denominado "problema de Hume" al problema de la inducción, tal terminología ha sido adoptada universalmente. He rebuscado inútilmente en la literatura sobre el tema intentando descubrir si alguien antes que yo había denominado "problema de Hume" al problema de la inducción. Todos los casos que he podido hallar pueden remitirse a escritores (tales como Russell o von Wright) que habían leído mi libro con mayor o menor atención. Evidentemente, es posible que haya pasado por alto algún autor anterior y nada hay más banal que reclamar prioridad en la introducción del nombre de un problema. Menciono el asunto por la sencilla razón de que se ha puesto de moda llamar "problema de Hume" a un problema totalmente distinto y algunos autores posteriores han intentado convencerme de que el "problema de la inducción de Hume" es, de hecho, distinto del que yo había bautizado con ese nombre.

Obviamente, hay varios problemas distintos que podría denominarse así. Mencionaré dos grupos [53]:

Grupo A. ¿Cómo podemos justificar la inducción?

Grupo B. ¿Se puede justificar la inducción? ¿Hay alguna razón para pensar que es justificable?

Como se verá inmediatamente, el grupo B plantea la pregunta fundamental: si se resuelve mediante una respuesta negativa clara, entonces no puede surgir la pregunta planteada en el grupo A.

Pretendo haber resuelto en este sentido la pregunta planteada en el grupo B. En otras palabras, pretendo haber resuelto el problema de la inducción de Hume en su versión más profunda. Lo digo explícitamente porque algunos filósofos han denominado "problema de la inducción de

---

[52] Esta sección (como tantas otras) se solapa en parte con el capítulo 1 de este libro. No obstante, la he mantenido porque creo que complementa dicho capítulo en diversos aspectos. (Véase también la nota 41, *supra*.)

[53] John Watkins llamó mi atención sobre un "Grupo C": ¿Es indispensable la inducción (i.e., algo basado en la repetición) sea o no justificable? "Lo que Hume suponía", dice Watkins, es que era indispensable. Es precisamente esto lo que niego, basándome en mi solución al problema de Hume. Lo único que precisamos suponer, en el mundo 3, es el realismo. En el mundo 2, nos vemos obligados a actuar, aceptando por regla general más de lo que es posible justificar, pero, con todo, elegimos la mejor de las diversas hipótesis rivales: es una consecuencia del realismo. Watkins piensa que el grupo C es el fundamental, pero no veo por qué habría de ser así. Aunque en cierto sentido es indispensable *elegir*, no ocurre lo mismo con la inducción. (Espero no haber entendido mal a Watkins.)

Hume" solamente al grupo A, atribuyéndome [54] erróneamente la tesis de que el problema de la inducción de Hume es *insoluble,* cuando lo que yo pretendo es haberlo resuelto completa aunque negativamente.

El problema de la inducción de Hume consta de dos elementos:

(a) El problema de justificar la validez de la pretensión de haber establecido con certeza (o con probabilidad, por lo menos) la verdad de una regla o generalización (o al menos su verdad probable) a partir de elementos de juicio singulares.

(b) La tesis de que la inducción está ligada a la *repetición* (y que la repetición está ligada a su vez al refuerzo de las asociaciones).

Naturalmente, podemos llamar inducción a lo que queremos. Podemos llamar teoría de la inducción a mi teoría de la crítica y el aumento de conocimiento. Sin embargo, creo que ello contribuiría poco a la claridad y mucho a la confusión. Esto es así porque el elemento (a) sobre si la inducción constituye una inferencia válida, es decir, suministra razones válidas en apoyo de la verdad de la proposición inducida, me parece característico del problema de Hume y de su respuesta (lógica) negativa; y (b) el elemento de repetición y asociación me parece característico del proplema de Hume, posibilitando la parte positiva (psicológica) de su respuesta.

La razón de todo esto estriba en que Hume resolvió los problemas planteados por (a) y (b) de dos modos especialmente distintos.

(a') Dijo que la inducción no era en absoluto válida como inferencia. No hay ni rastro de argumento lógico que apoye la inferencia de una generalización a partir de enunciados relativos al pasado (tales como repeticiones pasadas de algunos "elementos de juicio").

(b') Dijo también que a pesar de su carencia de validez lógica, la inducción desempeñaba una parte indispensable en la vida práctica. Vivimos confiando en repeticiones. La asociación reforzada mediante la repetición es el mecanismo fundamental de nuestro intelecto mediante el que vivimos y actuamos.

Por tanto, hay aquí una paradoja. *Ni siquiera nuestro intelecto funciona racionalmente. El hábito, racionalmente indefendible, es la fuerza fundamental que guía nuestro pensamiento y nuestras acciones.*

Esto hizo que Hume, uno de los pensadores más razonables de todos los tiempos, abandonase el racionalismo y considerase al hombre no como dotado de razón, sino como un producto del hábito ciego.

Según Russell esta paradoja de Hume es responsable de la esquizofrenia del hombre moderno. Esté o no en lo cierto Russell, sostengo que he hallado la solución.

La solución de la paradoja es que no sólo razonamos racionalmente y, por tanto, de modo contrario al principio de inducción cuya invalidez fue establecida por Hume, sino que además actuamos racionalmente de acuerdo con la razón más bien que con la inducción. Actuamos basándonos no en

---

[54] Véase la nota crítica de G. J. Warnock acerca de *L. Sc. D.* en *Mind,* New Series, *69,* 1960; pág. 100.

repetición o "hábito", sino en nuestras teorías mejor contrastadas que, como hemos visto, son las que se ven apoyadas por buenas razones racionales; naturalmente, poseemos buenas razones para creer no que son verdaderas, sino que son las *más válidas* desde el punto de vista de la búsqueda de la verdad o verosimilitud —las mejores entre las teorías rivales, las mejores aproximaciones a la verdad. Para Hume, el problema fundamental era: ¿Actuamos de acuerdo con la razón o no? Mi respuesta es: sí.

Con esto se resuelve la paradoja de Hume. Estaba en lo cierto con su crítica lógica a la posibilidad de una inducción válida. Lo que estaba equivocado era su psicología asociacionista, su creencia en que actuamos basándonos en el hábito concebido como resultado de una simple repetición.

Esta solución de la paradoja de Hume no afirma, naturalmente, que seamos criaturas enteramente racionales. Sólo afirma que en nuestra constitución humana no hay conflicto entre la racionalidad y la acción práctica.

Hay que añadir, evidentemente, que muchas veces la norma racional de nuestras acciones prácticas va muy por detrás de la norma aplicada a las fronteras del conocimiento: a menudo actuamos basándonos en teorías superadas hace mucho tiempo, en parte porque la mayoría de nosotros no entendemos lo que ocurre en las fronteras del conocimiento. No obstante, no creo que valga la pena seguir haciendo estas consideraciones.

30. Embrollos relacionados con el problema de la inducción

El propio Hume confundía el problema de la inducción con el problema de la conexión necesaria entre causa y efecto. Por otro lado, Kant vio en el problema de la validez *a priori* de la ley de la causalidad uno de los problemas más fundamentales de la metafísica. Hay que reconocer a Hume el mérito de haber formulado el problema lógico puro de la inducción y sus soluciones (estoy orgulloso de haber sido el primero, que yo sepa, en haber reconocido su mérito en este punto). Escribe, por ejemplo, que no poseemos *razones* para creer "*que los casos de los que no tenemos experiencia [deban de] asemejarse a aquellos de los que tenemos experiencia*" [55].

Esta formulación no podría distinguirse, de un modo más tajante, del problema de la necesidad causal que tan a menudo empaña la claridad del pensamiento de Hume. Esta formulación también se ve libre del elemento confuso de la inferencia del pasado al futuro. Lo único que se dice es que poseemos elementos de juicio en favor de la verdad de ciertos casos, afirmando que ello no nos autoriza a concluir o a extrapolar respecto a otras experiencias en otros casos (sean pasados o futuros).

---

[55] David Hume, *Treatise of Human Nature*, 1739, Parte III, sección VI; Selby-Bigge, pág. 89 [trad cit., pág. 150]. (El subrayado es del propio Hume.) Véase también mi *Logic of Scientific Discovery*, especialmente la pág. 369 [trad. cit., pág. 344], como ya dije en la nota 50.

Así pues, esto es, en toda su pureza, lo que he bautizado como "el problema [lógico] de la inducción de Hume".

La solución de Hume es todo lo clara que puede ser: no hay ningún argumento de la razón que autorice la inferencia de un caso a otro, por muy similares que puedan ser las condiciones. A este respecto, por lo demás, estoy plenamente de acuerdo con él.

No obstante, creo que Hume se equivoca cuando piensa que en la práctica hacemos tales inferencias basándonos en la repetición o hábito. Mantengo que es la suya una psicología primitiva [56]. Lo que hacemos en la práctica es saltar a la conclusión (frecuentemente a la manera del "troquelado" de Lorenz); es decir, a hipótesis muy poco convincentes a las que frecuentemente nos aferramos y con las que podemos perecer, a menos que seamos capaces de corregirlas, lo cual es especialmente factible si, a nivel humano, se formulan extrasomáticamente por escrito y se someten a crítica.

Afirmar que poseemos una inclinación irracional a dejarnos impresionar por el hábito y la repetición es muy distinto de afirmar que disponemos de un impulso que nos lleva a ingeniar hipótesis audaces que hemos de corregir si no queremos perecer. Lo primero describe un procedimiento de instrucción típicamente lamarckiano; lo segundo, un procedimiento de selección darwinista. Lo primero, como observó Hume, es irracional, mientras que lo segundo no parece contener ningún elemento irracional.

31. ¿Qué es lo que queda del problema erróneo de la justificación de la inducción?

El problema erróneo del grupo A —*el problema de la justificación de la inducción*— lo plantean quienes se dejan impresionar por la "Uniformidad de la Naturaleza": por el hecho de que el sol sale todos los días (una vez cada veinticuatro horas o cada 90.000 latidos del corazón); por el hecho de que todos los animales y hombres están abocados a la muerte [57]; y por el famoso ejemplo de Hume de que el pan alimenta. Sin embargo, estos tres ejemplos se refutan tomándolos en su sentido original [58].

"El sol sale todos los días" quería decir "vayas a donde vayas, el sol saldrá todos los días". Que era esto lo que se quería decir originalmente

---

[56] Puede haber otras psicologías tan malas como la de Hume, pero que no entran en conflicto con la lógica. Es más, afirmo que hay una psicología que, de hecho, está dominada por la lógica: la psicología racional del ensayo y supresión de errores.

[57] El término griego θνητός, traducido habitualmente por "mortal", significa de hecho "abocado a morir". Según esto, "Todos los hombres son mortales" se traduciría con mayor propiedad: "Todos los hombres están abocados a la muerte"; en este sentido, no se puede decir que el enunciado sea válido, ya que se deriva de "Todas las criaturas generadas están (por esencia) abocadas a la muerte", cosa que refutan las bacterias.

[58] He utilizado estos ejemplos con mucha frecuencia en mis conferencias, así como en el capítulo 1, págs. 23 y sigs. y nota 17. Sin embargo, he decidido dejar estos solapamientos para que ambos capítulos se puedan leer independientemente.

queda de manifiesto por el hecho de que Piteas de Marsella, el primer viajero conocido que atravesó el círculo polar y describió "el mar helado y el sol de medianoche", fue tomado durante siglos como paradigma de mentiroso y por el hecho de que la expresión "cuentos de viajes" surgió de ahí. Aristóteles dedujo el inevitable destino de todos los hombres que los lleva a la muerte del hecho de que todo lo generado —especialmente las criaturas vivas— debe corromperse, tesis que los biólogos ya no aceptan en absoluto de modo general (puesto que ya han conseguido mantener el corazón de un pollo *in vitro* latiendo más de medio siglo). Finalmente, el ejemplo de Hume de que el pan alimenta se vio básicamente refutado cuando el pan cocido de modo habitual barrió del mapa cierto pueblo francés a causa de un brote de ergotismo.

Pero, ¿todo se reduce a esto? Sí. Digan lo que digan los filósofos, es un hecho que estamos seguros por el sentido común de que mañana el sol saldrá sobre Londres. Sin embargo, no poseemos un conocimiento cierto de ello. Hay millones de posibilidades que pueden impedirlo. Quien intenta darnos razones positivas para creer en ello, no ha captado el problema. Hay que admitir que todos nosotros, seamos o no partidarios de Hume, deseamos que el sol continúe saliendo. Hay que admitir que es éste un deseo necesario —necesario para la acción, para la vida. Pero ni siquiera un deseo necesario constituye un conocimiento objetivo, aunque nos anime a creer.

En otras palabras, es falsa la confianza en esas reglas que aún esgrimen muchos filósofos como ejemplos típicos de reglas inductivas o fiables, a pesar de que parecen una excelente aproximación a la verdad.

Pero esto equivale solamente a mostrar la infiabilidad de la *llamada* inducción. Sin embargo, la genuina inducción por repetición no existe. Lo que se asemeja a la inducción es el razonamiento hipotético bien contrastado y corroborado de acuerdo con la razón y el sentido común, ya que hay un método de corroboración —el intento de refutar una teoría cuando parece posible una refutación. Si fracasa este intento, puede hacerse con fundamento racional la conjetura de que constituye una buena aproximación a la verdad— en todo caso, mejor que su predecesora.

¿Pero, acaso no podemos obtener algo así como la seguridad? ¿No podemos obtener seguridad con la inducción, con incontables casos de repetición?

La respuesta es negativa. (Es lo que dijo Hume.) Podemos obtener fácilmente la seguridad de sentido común —no tanto por repetición como por contrastaciones rigurosas—. Tengo tanta confianza como el que más en que el sol saldrá mañana sobre Londres o en que moriré dentro de poco, aunque de momento el pan siga alimentándome. Pero como teórico, se que pueden ocurrir otras cosas. Incluso se que el sol no sale diariamente en todas las partes de Europa, que las bacterias no mueren siempre, sino que se dividen y que el agua, el pan, el aire y nuestro entorno más ordinario y digno de confianza puede contener venenos mortales (y es de temer que pronto sea así).

También podemos preguntar: ¿por qué tenemos éxito en nuestra construcción de teorías? Respuesta: hasta ahora hemos tenido éxito, pero podemos fracasar mañana. Todo argumento que muestre que hemos de tener éxito probará demasiado. Lo único que podemos hacer es la conjetura de que vivimos en una parte del cosmos en que las condiciones para la vida y para el éxito de nuestra empresa cognoscitiva parecen ser favorables por el momento. Pero si conocemos algo, sabemos también que en casi todas las demás partes del cosmos son altamente desfavorables las condiciones para la vida y el conocimiento, ya que la cosmología nos enseña que el mundo está casi en todas partes completamente vacío y cuando no está vacío, está casi siempre ardiendo.

Por otro lado, el hecho de que los vehículos tirados por caballos fuesen durante muchos siglos algo cotidiano, no ha impedido su desaparición y sustitución por coches a motor. La aparente "uniformidad de la naturaleza" es muy poco digna de confianza; y aunque podamos decir que las leyes de la naturaleza no cambian, esto se aproxima peligrosamente a la afirmación de que en nuestro mundo hay ciertas conexiones abstractas que no cambian (lo cual es bastante trivial, si admitimos que no conocemos cuáles son, sino que hacemos conjeturas acerca de ellas) y que denominaremos "leyes de la naturaleza".

32. Escepticismo dinámico: confrontación con Hume

La posición sostenida aquí difiere radicalmente de lo que se ha dado en llamar "escepticismo" en la época moderna, al menos desde la Reforma, puesto que entonces se definía el escepticismo como la teoría pesimista respecto a la posibilidad de conocimiento. Sin embargo, el punto de vista expuesto aquí apoya esperanzadoramente la posibilidad del aumento de conocimiento *y, por tanto, del conocimiento.* Lo único que hace es eliminar el carácter de certidumbre que el sentido común consideraba esencial para el conocimiento, mostrando que tanto la certeza como el conocimiento son distintos de lo que presumía la teoría del sentido común. Difícilmente se podrá considerar escéptico a quien crea en la posibilidad del crecimiento ilimitado del conocimiento.

Por otra parte, algunos escépticos clásicos, como Cicerón y Sexto Empírico, no estaban muy lejos de la posición defendida aquí. Aunque sea extraño, se puede traducir perfectamente "*scepsis*" por "investigación crítica" y el "escepticismo dinámico" se podría identificar con la "investigación crítica potente" o incluso, a este respecto, con "investigación crítica esperanzadora", ya que por menguada que sea la esperanza, tiene un fundamento totalmente racional. Ciertamente, esto tiene poco que ver con el deseo de conocer cuando nada se puede conocer.

A este respecto, me parece importante volver a nuestro punto de partida —sentido común *más* razonamiento crítico— para recordar la conclusión de que el sentido común incluye el *realismo* —algo tal vez no muy

alejado del "realismo científico"— y que todos los argumentos conocidos en contra del realismo [59] resultan ser críticamente insostenibles o, más exactamente, desatinos insostenibles del aspecto más débil del sentido común, de la teoría del conocimiento del sentido común. Por tanto, no tenemos ninguna razón para abandonar el realismo.

Ahora bien, esto significa un cambio radical en la situación de mi "escepticismo esperanzador" especialmente cuando se compara con el de David Hume.

Hume arguye:

(1) La inducción (es decir, inducción por repetición) es totalmente inválida desde un punto de vista racional.

(2) De hecho, en nuestras acciones (y por tanto, en nuestras creencias) confiamos en la existencia de cierta realidad que no es totalmente caótica.

(3) A la vista de (1), esta confianza es irreparablemente irracional.

(4) La naturaleza humana es esencialmente irracional.

Acepto plenamente las tesis (1) y (2) de Hume, aunque rechazo la tesis (3) acerca de la irracionalidad. La rechazo porque no pretendo basar (2) en (1), sino que sostengo el realismo porque constituye una parte del sentido común que, hasta ahora, no se ha visto alcanzado por la crítica y que no tenemos razones para abandonar. En razón de su errónea teoría del conocimiento del sentido común, Hume creía que sólo podía ser razonable aceptar (2) cuando lo "conocemos" —es decir, cuando tenemos razones suficientes para creerlo—; por eso creía que dicha creencia se fundaba, *de facto,* en la inducción (que rechazaba con todo derecho por irracional). Pero no hay solamente un conocimiento humano de razón suficiente; también hay un conocimiento conjetural objetivo (y su análogo subjetivo, ya discutido en la sección 20). La condición de nuestra visión de sentido común acerca de la realidad no es esencialmente distinta [60] de las percepciones o impresiones inmediatas que Hume aceptaba como seguras: constituyen un conocimiento conjetural que forma parte de nuestro aparato orgánico presidido por el método de ensayo y eliminación de errores. Por tanto, no hay ninguna razón para fundamentar (2) en (1) o para considerar que necesita otro apoyo positivo que la ausencia de argumentos críticos sostenibles en contra suya.

Resumiendo, no necesitamos argumentar en favor del realismo partiendo de la inducción, como hizo Hume; no hay nada irracional en la conjetura del realismo; y los argumentos generales en contra suya, en cuya validez creía Hume, forman parte de su errónea esistemología de sentido común.

---

[59] Entre éstos no cuento el argumento válido en favor de un tipo de idealismo que no es incompatible con el realismo: el conocimiento humano es un producto humano, siendo todas las teorías inventos nuestros. Véase la nota 31, *supra,* así como *Conjectures and Refutations,* pág. 117 [trad. cit., pág. 139].

[60] En este punto estaba en lo cierto Thomas Reid. Véase, más arriba, el final de la sección 21.

Por tanto tenemos plena libertad para rechazar las tesis (3) y (4) de Hume.

Aún se puede añadir algo sobre (3) y (4). Creemos *esperanzadoramente* en el realismo, mas esa esperanza no es racional puesto que hay al menos algunos argumentos del "realismo científico" que nos obligan a predecir la destrucción final de toda vida.

Con todo, ni siquiera esto apoya las tesis (3) y (4) de Hume, pues no es irracional tener esperanza mientras vivimos —además estamos obligados constantemente a actuar y decidir.

33. Análisis de un argumento sacado de la improbabilidad de los accidentes

Como ya señalé brevemente (sección 22), creo que es un error considerar la probabilidad subjetiva como medida de la "*creencia racional*", pues no puede aportar nada positivo a la teoría del conocimiento.

Pero, puesto que nada depende de las palabras, no tengo nada en contra de que alguien llame conjetura "probable" (o la más probable de las conjeturas conocidas) a lo que yo he denominado aquí "buena" (o "la mejor") conjetura, a condición de que la palabra "probabilidad" no se interprete en el sentido del cálculo de probabilidades.

En mi opinión, la probabilidad en el sentido del cálculo de probabilidades no tiene nada que ver con la bondad de una hipótesis. (Como ya se ha explicado, sólo se puede utilizar la improbabilidad como medida de su contenido y, por tanto, de un aspecto de su bondad.)

Sin embargo, hay un viejo argumento, que posee cierta débil plausibilidad, que se puede poner en conexión con el cálculo de probabilidades del modo siguiente.

Supongamos que tenemos una hipótesis H que es lógicamente muy improbable; es decir, que tiene un contenido muy grande y hace afirmaciones en campos completamente desconexos hasta el momento. (Ejemplo: la teoría gravitatoria de Einstein no sólo predecía los movimientos planetarios de Newton, sino también una pequeña desviación en la órbita de Mercurio, un efecto sobre la trayectoria de los rayos de la luz al pasar por las proximidades de un cuerpo pesado y un corrimiento hacia el rojo de las líneas espectrales emitidas en campos gravitacionales fuertes.) Si todas estas predicciones se contrastan con éxito, entonces el siguiente argumento parece intuitivamente pertinente y razonable.

(1) *Difícilmente puede resultar accidental* el que la teoría prediga estas cosas tan manifiestamente improbables si no es verdadera. Partiendo de esto, se arguye que la probabilidad de su verdad es tan grande como la improbabilidad de que estos éxitos se deban a una acumulación de accidentes.

No creo que el argumento (1), en la forma planteada, pueda tomarse

como perfectamente válido, aunque no obstante creo que en él hay algo que merece la pena. Examinémoslo con más detalle.

Supongamos que el argumento (1) es válido. Entonces podríamos estimar que *la probabilidad de que la teoría sea verdadera* es, al menos, la probabilidad de que sea verificada de un modo puramente accidental; y, si los efectos predichos son lógicamente muy improbables —por ejemplo, porque las cantidades numéricas se predicen con toda precisión y corrección— entonces, el producto de estos números tan pequeños será el número a deducir de la unidad. En otras palabras, con este método de cálculo obtendremos una probabilidad muy próxima a la verdad para conjeturas buenas [61].

El argumento parece convincente a primera vista, pero es claramente inválido. Tomemos la teoría de Newton (N). Arroja predicciones tan precisas que según nuestro argumento daría una probabilidad muy próxima a la unidad. La teoría de Einstein (E) arrojaría una probabilidad aún mayor. Pero, según el cálculo de probabilidades, tenemos (escribiendo "V" en lugar de "o"):

$$p(N \vee E) = p(N) + p(E) - p(NE);$$

y puesto que las teorías son incompatibles, de manera que $p(NE) = 0$, obtenemos

$$p(N \vee E) = p(N) + p(E) \approx 2$$

(es decir, casi 2) lo cual es absurdo.

La solución del problema consiste en mostrar que el razonamiento (1) es especioso, pues se puede decir lo siguiente.

(2) El buen acuerdo con los improbables resultados observados ni es un accidente ni se debe a la verdad de la teoría, sino tan sólo a su *verosimilitud*.

Este argumento (2) explicará por qué muchas teorías incompatibles pueden coincidir en muchos aspectos refinados, en los que intuitivamente sería altamente improbable que coincidiesen por puro accidente [62].

Por tanto, el argumento (1) se puede exponer un poco más correctamente como sigue.

(1') Hay verosimilitud, razón por la cual el acuerdo entre teoría y hecho, muy improbable por accidente, se puede interpretar como índice de que la teoría posee (comparativamente) un alto grado de verosimilitud. Hablando en general, un mayor acuerdo en puntos improbables se puede interpretar como índice de mayor verosimilitud.

No creo que se puedan esgrimir muchas cosas en contra de este argumento, aunque me disgustaría que se desarrollase hasta convertirse en otra

---

[61] Se trata de un viejo argumento expuesto de modo ligeramente distinto. Se pueden encontrar trazas suyas en la *Etica a Nicómaco* de Aristóteles y en el *Liber de Astronómica* de Teón de Esmirna, ed. Th. H. Martin, París, 1949, pág. 293.

[62] No estoy seguro de haber publicado nunca este argumento, pero recuerdo haberlo tenido en cuenta por primera vez alrededor del año 1930.

teoría más de la inducción. Pero quiero dejar muy claro que el grado de corroboración de una teoría (que es algo así como una medida del rigor de las contrastaciones que ha pasado) no se puede interpretar sencillamente como medida de su verosimilitud. A lo sumo, sólo es un *indicador* (como expliqué en 1960 y 1963 cuando introduje por primera vez la idea de verosimilitud; véase por ejemplo *Conjectures and Refutations*, págs. 234 y siguientes. [Traducción castellana, págs 272 y sigs.]) de verosimilitud, tal como aparece en un momento *t*. He introducido el término "corroboración" para el grado en que una teoría ha sido rigurosamente contrastada. Se emplea fundamentalmente a efectos de comparación: por ejemplo, *E* está más rigurosamente contrastada que *N*. El grado de corroboración de una teoría tiene siempre un índice temporal: es el grado en que una teoría aparece como bien contrastada en el momento *t*. Esto no puede constituir una medida de su verosimilitud, sino que ha de tomarse como índice de cómo *aparece* su verosimilitud en el momento *t* comparada con otra teoría. El grado de corroboración es, pues, una guía de la preferencia entre dos teorías en un cierto estadio de la discusión respecto a su aparente aproximación a la verdad en ese momento. Ahora bien, lo único que nos dice es que una de las teorías ofrecidas *parece —a la luz de la discusión—* la más próxima a la verdad.

34. Sumario: Filosofía crítica del sentido común

Una vez que hemos visto la necesidad de una filosofía crítica, surge el problema del punto de partida. ¿Dónde hemos de empezar? La cuestión parece importante porque hay el peligro de que un error inicial pueda acarrear las peores consecuencias.

Respecto a este punto de partida difieren radicalmente los puntos de vista sostenidos por la mayoría de los filósofos clásicos y contemporáneos por un lado y, por otro, los que he propuesto aquí a modo de semi-filosofía del sentido común. Voy a intentar resumir en una tabla las diferencias más importantes.

Filósofos anteriores	*Mi enfoque crítico*
(1) La elección de nuestro punto de partda es decisivamente importante: hemos de cuidarnos de no incurrir en un error en los comienzos mismos.	(1') La elección de nuestro punto de partida no es decisivamente importante porque se puede criticar y corregir como todo lo demás.
(2) En la medida de lo posible, nuestro punto de partida debería ser verdadero y cierto.	(2') No hay manera de encontrar un punto de partida con esas características.

(3) Se puede encontrar en la experiencia personal del yo (subjetivismo) o en la descripción pura de la conducta (objetivista) [68].

(3') Puesto que no se puede encontrar ni en el subjetivismo ni en el objetivismo, lo mejor será partir de ambos para criticarlos.

(4) Al aceptar este tipo de subjetivismo o de objetivismo, los filósofos aceptan acríticamente *una forma de la teoría del conocimiento de sentido común,* teoría que puede considerarse como el punto más débil del sentido común.

(4') Siempre que seamos críticos con todo lo que se diga en nombre del sentido común, es recomendable partir de él, por muy vagos que sean sus puntos de vista.

(5) La teoría que aceptan los subjetivistas dice que el conocimiento más cierto que nos es dado poseer versa sobre nosotros mismos y nuestras experiencias observacionales o perceptivas. (Los objetivistas y subjetivistas coinciden en hacer incapié en la certeza de las experiencias perceptivas.)

(5') Una pequeña reflesión crítica nos convence de que todo nuestro conocimiento está impregnado de teoría y en su mayoría posee también un carácter conjetural.

(6) Hay algunos hechos sólidos sobre los cuales se puede construir el conocimiento, tales como nuestros datos de los sentidos o sensaciones claras y distintas: las experiencias directas e inmediatas no pueden ser falsas.

(6') Puesto que todo conocimiento está impregnado de teoría, está construido sobre arena; pero puede mejorarse dragando críticamente en profundidad y no dando por supuesto ningún sediciente "dato".

(7) Esto constituye un resultado evidente de la teoría del conocimiento del sentido común.

(7') Es aquí donde falla la teoría del conocimiento del sentido común: pasa por alto el carácter indirecto y conjetural del conocimiento. Incluso nuestros órganos de los sentidos (para no hablar de la interpretación de sus dictámenes) están impregnados de teoría y sujetos a error, aunque ésto ocurra en contadas ocasiones en los organismos sanos.

(8) Pero la teoría del conocimiento del sentido común, que siempre parte de una forma del realismo,

(8') Reconocemos que tanto el realismo como su teoría (biológica) del conocimiento son dos conjetu-

[68] Este tipo de objetivismo recibe ordinariamente el nombre de "conductismo" u "operacionalismo". No se discute detalladamente en este escrito.

acaba hundiéndose en la ciénaga del idealismo u operacionalismo epistemológico.

(9) El sentido común, habiendo partido del realismo para terminar en el subjetivismo, se refuta a sí mismo. (Puede decirse que ésto forma parte del punto de vista de Kant.)

ras, y argüímos que la primera es una conjetura muy superior al idealismo.

(9') La teoría del conocimiento del sentido común se refuta por auto-contradictoria; pero ésto no afecta a la teoría del sentido común respecto al mundo; es decir, al realismo.

Todo intento de mantener la teoría del sentido común como un todo íntegro —realismo más epistemología del sentido común— está abocado al fracaso. Así, mediante el escepticismo relativo al propio punto de partida, la teoría del sentido común se escinde en dos partes como mínimo —realismo y epistemología— y podemos rechazar esta última, sustituyéndola por una teoría objetiva que utilice la primera.

3. EPISTEMOLOGIA SIN SUJETO COGNOSCENTE*

Permítaseme comenzar con una confesión. Aunque soy un filósofo muy afortunado, tras toda una vida como conferenciante no me hago ilusiones acerca de lo que pueda comunicar en una charla, razón por la cual no intentaré convencerles. En vez de ello, procuraré lanzarles un desafío y, si es posible, provocarles.

1. TRES TESIS SOBRE EPISTEMOLOGÍA Y EL TERCER MUNDO

Podría haber lanzado un reto a quienes han oído hablar de mi actitud hostil hacia Platón y Hegel titulando esta conferencia "*Teoría del Mundo Platónico*" o "*Teoría del Espíritu Objetivo*".

El tema central de esta conferencia será lo que acostumbro a llamar, a falta de un nombre mejor, "*el tercer mundo*". Para explicar esta expresión, habrá que señalar que, sin tomar demasiado en serio las palabras "mundo" o "universo", podemos distinguir los tres mundos o universos siguientes: primero, el mundo de los objetos físicos o de los estados físicos; en segundo lugar, el mundo de los estados de conciencia o de los estados mentales o, quizá, de las disposiciones comportamentales a la acción; y en tercer lugar, el mundo de los *contenidos de pensamiento objetivo*, especialmente, de los pensamientos científicos y poéticos y de las obras de arte.

Así pues, es obvio que lo que denomino "el tercer mundo" tiene mucho que ver con la teoría de las Formas o Ideas de Platón y, por tanto, también con el espíritu objetivo de Hegel, si bien mi teoría difiere radicalmente en algunos aspectos decisivos de las de Platón y Hegel. Tiene aún más que ver con la teoría de Bolzano sobre el universo de las proposiciones en sí mismas y de las verdades en sí mismas, aunque también difiere de la de Bolzano. Mi tercer mundo se asemeja, en mayor medida, al universo de los contenidos objetivos del pensamiento de Frege.

Mi punto de vista o mi razonamiento no entraña que no podamos hacer otras clasificaciones de nuestros mundos o que no podamos prescindir absolutamente de cualquier clasificación. Especialmente, podríamos distinguir más de tres mundos. El término "tercer mundo" no es más que una cuestión de conveniencia.

---

* Ponencia presentada el 25 de agosto de 1967 al Tercer Congreso Internacional de Lógica, Metodología y Filosofía de la Ciencia, 25 de agosto-2 de septiembre de 1967; publicada por primera vez en las actas de este Congreso, eds. B. van Roosetlsaar y J. F. Staal. Amsterdam 1968, págs. 333-73.

Al defender un tercer mundo objetivo deseo provocar a aquellos que yo llamo "*filósofos de la creencia*": los que, como Descartes, Locke, Berkeley, Kant o Russell, se interesan por nuestras creencias subjetivas y su fundamento u origen. Frente a ellos subrayo que el problema consiste en encontrar teorías mejores y más audaces y que lo importante *no* son las *creencias,* sino la *preferencia crítica.*

No obstante, deseo confesar desde ahora mismo que soy realista: sugiero, un poco como el realismo ingenuo, que hay mundos físicos y un mundo de estados de conciencia y que ambos interactúan entre sí. También creo que hay un tercer mundo en un sentido que voy a explicar más a fondo.

Entre los inquilinos de mi "tercer mundo" se encuentran especialmente los *sistemas teóricos* y tan importantes como ellos son los *problemas* y las *situaciones problemáticas.* Demostraré también que los inquilinos más importantes de este mundo son los *argumentos críticos* y lo que podríamos llamar —por semejanza con los estados físicos o los estados de conciencia— *el estado de una discusión* o el *estado de un argumento crítico,* así como los contenidos de las revistas, libros y bibliotecas.

Naturalmente, la mayoría de los que se oponen a la tesis del tercer mundo objetivo admiten que hay problemas, conjeturas, teorías, argumentos, revistas y libros. Mas todos ellos suelen decir que estas entidades son esencialmente *expresiones* simbólicas o lingüísticas de estados mentales subjetivos o, tal vez, de disposiciones comportamentales a la acción. Dicen, además, que estas entidades son medios de *comunicación* —es decir, medios simbólicos o lingüísticos mediante los que despertamos en los demás estados mentales o disposiciones comportamentales a la acción semejantes a los nuestros.

Contra esto he objetado muy a menudo que no se pueden relegar todas estas entidades y su contenido al segundo mundo.

Repetiré uno de mis argumentos típicos [1] en favor de *la existencia* (más o menos) *independiente del tercer mundo.*

Examinaré dos experimentos mentales:

Experimento (1). Todas las máquinas y herramientas han sido destruidas, junto con todo nuestro aprendizaje subjetivo, incluyendo el conocimiento subjetivo sobre las máquinas, las herramientas y cómo usarlas. Sin embargo, sobreviven *las bibliotecas y nuestra capacidad de aprender en ellas.* Está claro que, tras muchas penalidades, nuestro mundo puede echar a andar de nuevo.

Experimento (2). Como antes, han sido destruidas las máquinas y herramientas, junto con nuestro aprendizaje subjetivo que incluye nuestro conocimiento subjetivo de las máquinas, las herramientas y cómo usarlas. Pero, esta vez, *también han sido destruidas todas las bibliotecas,* de manera que nuestra capacidad para aprender de los libros se hace inútil.

---

[1] El argumento es una adaptación de Popper 1962, vol. II; cf. pág. 108. (Las referencias o detalles bibliográficos de este capítulo se dan en la bibliografía selecta que aparece al final del capítulo. [Véase allí la traducción del libro mencionado, pág. 307.])

Si reflexionamos acerca de estos dos experimentos, tal vez se vea un poco más clara la realidad, significación y grado de autonomía del tercer mundo (así como sus efectos sobre el segundo y el primero). Esto es así porque, en el segundo caso, nuestra civilización no volvería a emerger hasta al cabo de muchos milenios.

Es mi deseo defender en esta conferencia tres tesis fundamentales que conciernen a la epistemología. Entiendo la epistemología como la teoría del *conocimiento científico*.

He aquí mi primera tesis. La epistemología tradicional ha estudiado el conocimiento o el pensamiento en un sentido subjetivo —en el sentido de la utilización ordinaria de las palabras "sé" ["*know*"] o "pienso". Sostengo que esto ha descarriado a los estudiosos de la epistemología: aunque su intención era estudiar el conocimiento científico, de hecho estudiaron cosas irrelevantes para el conocimiento científico, por la sencilla razón de que el *conocimiento científico* no es el conocimiento tomado en el sentido de la utilización ordinaria de la palabra "sé". Mientras que el conocimiento en el sentido de "sé" pertenece a lo que denomino "segundo mundo" el mundo de los *sujetos*, el conocimiento científico pertenece al tercer mundo, al mundo de las teorías objetivas, de los problemas objetivos y de los argumentos objetivos.

Mi primera tesis afirma, por tanto, que la epistemología tradicional de Locke, Berkeley, Hume e incluso Russell es irrelevante en un sentido bastante estricto de la palabra. Esta tesis tiene como corolario que también es irrelevante una gran parte de la epistemología contemporánea. Aquí hay que incluir la moderna lógica epistémica, *suponiendo* que pretenda ser una teoría del *conocimiento científico*. Sin embargo, un lógico epistémico puede fácilmente inmunizarse de un modo completo contra mi crítica, por el sencillo expediente de dejar claro que no pretende contribuir a la *teoría del conocimiento científico*.

Mi primera tesis entraña la existencia de dos sentidos distintos de conocimiento o pensamiento: (1) *conocimiento o pensamiento en sentido subjetivo* que consiste en un estado mental o de conciencia, en una disposición a comportarse o a reaccionar y (2) *conocimiento o pensamiento en sentido objetivo* que consiste en problemas, teorías y argumentos en cuanto tales. El conocimiento en este sentido objetivo es totalmente independiente de las pretensiones de conocimiento de un sujeto; también es independiente de su creencia o disposición a asentir o actuar. El conocimiento en sentido objetivo es *conocimiento sin conocedor*: *es conocimiento sin sujeto cognoscente*.

Escribía Frege del conocimiento en sentido objetivo: "Entiendo por *pensamiento*, no el acto subjetivo de pensar, sino su *contenido objetivo*"... [2]

Los dos sentidos de pensamiento y sus interesantes interrelaciones se

[2] Cf. Frege, 1892, pág. 32; es mío el subrayado. [Cf. la traducción de Ulises Moulines, pág. 58, nota 5.]

pueden ilustrar mediante la siguiente cita, muy convincente, de Heyting (1962, pág. 195), que alude al acto mediante el cual Brouwer inventó su teoría del continuo.

"Si las funciones recursivas hubiesen sido inventadas anteriormente, [Brouwer] tal vez no hubiese elaborado la noción de secuencia de elección lo cual, según creo, hubiera sido una desgracia."

Esta cita hace referencia, por un lado, a ciertos *procesos subjetivos del pensamiento* de Brouwer, diciendo que, si la *situación problemática objetiva* hubiese sido diferente, no hubieran podido tener lugar (lo cual hubiese sido lamentable). Así pues, Heyting menciona ciertas *influencias* posibles sobre los procesos de pensamiento subjetivos de Brouwer, expresando su opinión acerca de su valor. Ahora bien, es interesante que las influencias, *qua* influencias, deban ser subjetivas: sólo la familiaridad subjetiva de Brouwer con las funciones recursivas hubiese podido tener el desafortunado resultado de impedirle inventar las secuencias de elección libre.

Por otro lado, la cita de Heyting señala cierta relación objetiva entre los *contenidos objetivos* de dos pensamientos o teorías: Heyting no se refiere a las condiciones subjetivas o a la electroquímica de los procesos cerebrales de Brouwer, sino a una *situación problemática objetiva en matemáticas* y su posible influencia sobre los actos de pensamiento subjetivos de Brouwer que se doblegaron ante la resolución de estos problemas objetivos. Expresaría esto diciendo que las consideraciones de Heyting versan sobre la *lógica situacional* objetiva o del tercer mundo del invento de Brouwer y que implican que la situación en el tercer mundo puede afectar al segundo. De un modo similar, la sugerencia de Heyting de que hubiese sido una desgracia que Brouwer no hubiese inventado las secuencias de elección, es una manera de decir que el *contenido objetivo* del pensamiento de Brouwer era valioso e interesante; es decir, valioso e interesante por el modo en que transformó la situación problemática objetiva del tercer mundo.

Simplificando, cuando digo que "el pensamiento de Brouwer recibió la influencia de Kant", o incluso que "Brouwer rechazó la teoría del espacio de Kant", estoy hablando, al menos en parte, de actos de pensamiento en sentido subjetivo: la palabra "influencia" señala un contexto de proceso o actos de pensamiento. Cuando digo, por el contrario, que "el pensamiento de Brouwer se diferencia en gran medida del de Kant", entonces está bastante claro que me refiero fundamentalmente a contenidos. Finalmente, cuando digo que "los pensamientos de Brouwer son incompatibles con los de Russell", al utilizar un *término lógico* como "*incompatible*", está perfectamente claro que utilizo la palabra "pensamiento" únicamente en el sentido objetivo de Frege, refiriéndome sólo al contenido objetivo, lógico, de las teorías.

Del mismo modo que el lenguajee ordinario no posee, desgraciadamente, términos distintos para "pensamiento" en el sentido del segundo y del tercer mundo, tampoco los posee para los dos sentidos correspondientes de "sé" y "conocimiento".

Mencionaré en primer lugar tres ejemplos subjetivos o del segundo mundo, en orden a mostrar la existencia de ambos sentidos:

(1) "*Sé* que usted intenta provocarme, pero no lo logrará."
(2) "*Sé* que el último teorema de Fermat no ha sido demostrado, pero creo que algún día lo será."
(3) "Según el artículo 'Conocimiento' de *The Oxford English Dictionary*, el *conocimiento* es un "estado que consiste en ser consciente o estar informado".

A continuación pondré tres ejemplos objetivos o del tercer mundo:

(1) Según el artículo "Conocimiento" de *The Oxford English Dictionary*, el *conocimiento* es una "parte del aprendizaje; una ciencia; un arte".
(2) "Teniendo en cuenta el estado actual del *conocimiento matemático*, parece posible que el último teorema de Fermat sea indecidible."
(3) "Aseguro que esta tesis constituye una *contribución* original y significativa *al conocimiento*."

Estos ejemplos tan triviales tienen por única función tratar de ayudar a clarificar lo que quiero decir cuando hablo de "conocimiento en sentido objetivo". Mi cita de *The Oxford English Dictionary* no ha de interpretarse, sea como una concesión al análisis lingüístico, sea como un intento de apaciguar a sus adherentes. No lo cito para mostrar que la "usanza ordinaria" recoge el "conocimiento" en el sentido objetivo de mi tercer mundo. De hecho, me sorprendió encontrar ejemplos de usanzas objetivas de "conocimiento" en *The Oxford English Dictionary*. (Me sorprendió aún más encontrar algunas usanzas objetivas, al menos *en parte*, de "conocer": "discernir... estar familiarizado con [una cosa, un lugar, una persona];... comprender". Por lo que sigue se verá que estas usanzas pueden ser objetivas en parte.) [3] En cualquier caso, mis ejemplos no pretenden ser argumentos, sino tan sólo ilustraciones.

Mi *primera tesis*, que hasta aquí no ha sido defendida con argumentos, sino tan sólo ilustrada, consistía en afirmar el carácter irrelevante de la epistemología tradicional —con su concentración en el segundo mundo— para el estudio del conocimiento científico.

Mi *segunda tesis* consiste en afirmar que lo que es relevante para la epistemología es el estudio de los problemas científicos objetivos y de las situaciones problemáticas, de las conjeturas científicas (que considero un modo más de referirse a las hipótesis o teorías científicas), de las discusiones científicas, de los argumentos críticos y del papel desempeñado por los elementos de juicio en los argumentos y, por tanto, de las revistas o libros científicos y de los experimentos y su valoración en argumentos científicos; resumiendo, que el estudio del tercer mundo del conocimiento objetivo, *en*

[3] Véase la sección 7.1, más adelante.

gran medida autónomo, es de importancia decisiva para la epistemología.

Un estudio epistemológico como el descrito en mi segunda tesis muestra que muy a menudo los científicos no pretenden que sus conjeturas sean verdaderas. Tampoco pretenden "conocerlas" en el sentido subjetivo de "conocer", ni creer en ellas. Aún cuando, en general, no pretendan conocer, al desarrollar sus programas de investigación actúan sobre la base de sospechas acerca de lo que es y no es fructífero y de qué línea de investigación promete más resultados en el tercer mundo del conocimiento objetivo. En otras palabras, los científicos actúan basándose en una sospecha o, si se quiere, en una *creencia subjetiva* (ya que podemos dar ese nombre a la base subjetiva de la acción) acerca de lo que es prometedor para el *desarrollo del conocimiento objetivo en el tercer mundo.*

Sugiero que esto constituye un argumento en favor tanto de mi *primera tesis* (sobre la irrelevancia de una epistemología subjetivista), como de mi *segunda tesis* (sobre la relevancia de una epistemología objetivista).

Pero tengo una *tercera tesis.* Es la siguiente: una epistemología objetivista que estudie el tercer mundo puede contribuir a arrojar muchísima luz sobre el segundo mundo de la conciencia subjetiva; especialmente, sobre los procesos de pensamiento subjetivos de los científicos. Pero *la conversa no es verdadera.*

Estas son mis tres tesis fundamentales.

Además de mis tres tesis fundamentales, ofrezco tres tesis de apoyo.

La primera de ellas afirma que el tercer mundo es un producto natural del animal humano, comparable a una tela de araña.

La segunda tesis de apoyo (que, por cierto, es una tesis casi crucial) afirma que el tercer mundo es *autónomo* en gran medida, aunque actuemos constantemene sobre él y éste, a su vez, actúe sobre nosotros: es autónomo a pesar de ser un producto nuestro y de tener un fuerte efecto de retroalimentación sobre nosotros; es decir, sobre nosotros en cuanto inquilinos del segundo e incluso del primer mundo.

La tercera tesis de apoyo afirma que el conocimiento se desarrolla mediante esta interacción entre nosotros y el tercer mundo, existiendo una estrecha analogía entre el crecimiento del conocimiento y el crecimiento biológico; es decir, la evolución de animales y plantas.

2. Aproximación biológica al tercer mundo

En la presente sección de mi charla intentaré defender la existencia de un mundo autónomo mediante una especie de argumento biológico o evolucionista.

Un biólogo puede interesarse por el comportamiento de los animales, pero ha de interesarse también por algunas de las *estructuras no-vivas* que producen los animales, como las telas de araña, los tejidos o nidos construidos por las avispas o las hormigas, las madrigueras de los tejones, los diques

construidos por los castores o las sendas que hacen los animales en la selva.

Distinguiré dos grandes tipos de problemas que surgen del estudio de estas estructuras. El primer tipo consta de problemas relativos a *los métodos empleados* por los animales o a *los modos de comportamiento de los animales* cuando construyen estas estructuras. Esta primera categoría consta *de problemas relativos a los actos de producción*: de disposiciones comportamentales del animal y de las relaciones entre el animal y el producto. La segunda categoría de problemas se ocupa de las *estructuras mismas*. Se ocupa de la química de los materiales empleados en la estructura, de sus propiedades geométricas y físicas, de sus cambios según las condiciones especiales del medio y de su dependencia y ajuste con esas condiciones ambientales. También es *muy* importante la *relación de retroalimentación* entre las propiedades de la estructura y el comportamiento de los animales. Al tratar esta segunda categoría de problemas —es decir, las estructuras mismas— hemos de considerar también las estructuras desde el punto de vista de sus *funciones* biológicas. Es así obvio que al discutir problemas de la segunda categoría surgirán problemas de la primera, por ejemplo: "¿Cómo se construyó este nido?" y "¿Qué aspectos de su estructura son típicos (y por tanto presumiblemente tradicionales o heredados) y cuáles son variaciones ante condiciones especiales?".

Como muestra el problema expuesto en el segundo ejemplo, los problemas de la primera categoría —es decir, los problemas relativos a la producción de la estructura— son sugeridos frecuentemente por los problemas de la segunda categoría. Debe ser así, ya que ambos tipos de problemas dependen del *hecho de que existan tales esrtucturas objetivas*, hecho que por sí mismo pertenece a la segunda categoría. Así, puede decirse que la existencia de *estructuras en sí mismas* crea ambas categorías de problemas. Podemos decir que la segunda categoría de problemas —problemas relacionados con las estructuras mismas— es más fundamental: lo único que presupone de la primera categoría es el simple hecho de que las esructuras son *producidas* de algún modo *por* algunos animales.

Ahora bien, como es natural, estas sencillas consideraciones pueden aplicarse también a los productos de la actividad *humana*, tales como casas o herramientas, y también a las obras de arte. Se aplican, lo que es de importancia especial para nosotros, a los que llamamos "lenguaje" y a lo que llamamos "ciencia" [4].

La conexión entre estas consideraciones biológicas y el tema de mi conferencia se pueden poner en claro reformulando mis tres tesis fundamenles. Todas ellas pueden formularse diciendo que en la actual situación problemática de la filosofía, pocas cosas son tan importantes como la conciencia de la distinción entre las dos categorías de problemas —problemas de producción por un lado, problemas relativos a las estructuras mismas producidas por el otro. Mi segunda tesis afirma que hemos de constatar que la segunda categoría de problemas, los relativos a los productos mismos, es

[4] Sobre estos "artefactos", cf. Hayek, 1967, pág. 111.

casi en todos los sentidos más importante que la primera— los problemas de producción. Mi tercera tesis dice que los problemas de la segunda categoría son básicos para comprender los problemas de producción: en contra de lo que pueda parecer a primera vista, estudiando los productos mismos podemos aprender sobre el comportamiento productivo más de lo que podemos aprender, sobre los productos, estudiando el comportamiento productivo. Esta tercera tesis puede considerarse una tesis anti-comportamentalista y anti-psicologista.

En su aplicación a lo que puede denominarse "conocimiento", mis tres tesis pueden formularse del modo siguiente:

(1) Deberíamos tener siempre en cuenta la distinción que hay entre los problemas relacionados con nuestras contribuciones personales a la producción del conocimiento científico, por una parte, y los problemas relacionados con la estructura de los diversos productos, como teorías o argumentos científicos, por la otra.

(2) Deberíamos constatar que el estudio de los productos es mucho más importante que el estudio de la producción, incluso para comprender la producción y sus métodos.

(3) Podemos aprender más sobre la heurística y la metodología e incluso sobre la psicología de la investigación estudiando las teorías y los argumentos en pro y en contra que empleando un método directo conductista, psicológico o sociológico. En general, podemos aprender muchísimo sobre el comportamiento o la psicología mediante el estudio de los productos.

De ahora en adelante llamaré "objetivo" o "propio del tercer mundo" al enfoque desde el punto de vista de los productos —teorías y argumentos— y, "subjetivo" o "propio del segundo mundo" al enfoque del conocimiento científico desde el punto de vista conductista, psicológico o sociológico.

El atractivo del enfoque subjetivo se debe en gran medida al hecho de que es *causal*, ya que admito que las estructuras objetivas, cuya prioridad proclamo, son causadas por la conducta humana. Al ser causal, el enfoque subjetivo puede parecer más científico que el objetivo, el cual, como si dijésemos, parte de los efectos más bien que de las causas.

Sostengo que se trata de un argumento equivocado, aunque admito que las estructuras objetivas son productos de la conducta. En todas las ciencias se procede de los efectos a las causas. El efecto plantea el problema —el problema a explicar, el explicandum— y el científico intenta resolverlo construyendo una hipótesis explicativa.

Mis tres tesis fundamentales, con su énfasis en el producto objetivo, no son, por tanto, ni teleológicas ni acientíficas.

3. La objetividad y autonomía del tercer mundo

Una de las razones fundamentales que explican el erróneo enfoque subjetivo del conocimiento es la sensación de que un libro no es nada sin un lector: sólo si es entendido, es realmente un libro; de lo contrario, no es más que un papel con manchas negras.

Este punto de vista es erróneo en muchos aspectos. Un nido de avispas sigue siendo un nido de avispas aún después de abandonado, aunque no se utilice ya nunca más como nido. Un nido de pájaros sigue siéndolo aunque nunca haya sido habitado. De un modo semejante, un libro sigue siendo tal —cierto tipo de producto—, aunque no se lea nunca (como puede ocurrir perfectamene hoy día).

Además, no es preciso que un libro, o incluso toda una biblioteca, haya sido escrito por alguien: una computadora, por ejemplo, puede producir e imprimir una serie de tablas de logaritmos. Incluso pueden ser las mejores tablas de logaritmos —pueden contener logaritmos hasta, digamos, cincuenta cifras decimales—. Pueden ser enviadas a las bibliotecas, aunque resulten demasiado engorrosas de manejar; en cualquier caso, pueden pasar años antes que alguien las utilice, y muchas de sus cifras (que representan teoremas matemáticos) puede que nunca sean consultadas mientras viva el hombre sobre la tierra. Sin embargo, esos números contienen lo que yo llamo "conocimiento objetivo"; no me preocupa el problema de si puedo o no llamarlo así.

Aunque el ejemplo de estas tablas de logaritmos puede parecer forzado, no lo es. Yo diría que casi todos los libros tienen las siguientes características: contienen conocimiento objetivo verdadero o falso, útil o inútil, siendo casi accidental que alguien los lea alguna vez y capte sus contenidos. La persona que lee un libro comprendiéndolo es una criatura excepcional. Pero, aunque fuese más común, siempre habría una gran cantidad de incompresiones y malas interpretaciones. Pero lo que convierte las manchas negras sobre papel blanco en un libro o en conocimiento en sentido objetivo no es el hecho, un tanto accidental, de evitar tales incomprensiones, sino algo más abstracto. Lo que hace de algo un libro es la posibilidad o potencialidad de ser comprendido, su carácter disposicional de ser comprendido e interpretado o incomprendido y mal interpretado. Ahora bien, esta potencialidad o disposición puede existir incluso sin ser actualizada o realizada nunca.

Esto se puede ver más claramente imaginando que unos sucesores nuestros civilizados encuentran algunos libros o bibliotecas tras la desaparición de la raza humana (no importa si nuestros sucesores son animales terrestres que han logrado civilizarse o si se trata de visitantes del espacio exterior). Estos libros pueden ser descifrados. Pueden ser aquellas tablas de logaritmos que nunca habían sido leídas anteriormente, según argumentábamos. Así pues, queda bien claro que para que algo sea un libro basta con que pueda ser descifrado, no siendo esencial ni que haya sido realizado por animales pensantes, ni que haya sido efectivamente leído o comprendido.

Por tanto, concedo que para que un libro pertenezca al tercer mundo del conocimiento objetivo ha de ser posible —en principio o virtualmente— que alguien lo comprenda (o lo descifre, lo entienda lo "conozca"). No concedo más.

Podemos decir, pues, que hay una especie de tercer mundo platónico (o bolzanesco) de libros en sí mismo, teorías en sí mismas, problemas en sí mismos, situaciones problemáticas en sí mismas, argumentos en sí mismos, etcétera. Afirmo, además, que aún cuando este tercer mundo sea un producto humano, hay muchas teorías, argumentos y situaciones problemáticas en sí mismos que nunca han sido producidos o entendidos por el hombre y puede que nunca lo sean.

La tesis de la existencia de este tercer mundo de situaciones problemáticas sorprenderá a muchos por su carácter excesivamente metafísico y dudoso, pero se puede defender poniendo de manifiesto una analogía biológica. Hay una analogía total, por ejemplo, en el terreno de los nidos de pájaro. Hace algunos años me hicieron un regalo para el jardín —una caja de nidificación para pájaros. Se trataba de un producto humano, evidentemente, no de los pájaros—, del mismo modo que nuestra tabla de logaritmos era un producto de la computadora más que del hombre. Pero en el contexto del mundo de los pájaros, formaba parte de una situación problemática efectiva, constituyendo una oportunidad objetiva. Durante muchos años, los pájaros ni siquiera parecieron percatarse de la presencia de la caja de nidificación. Pero luego, algunos paros azules la inspeccionaron cuidadosamente e incluso empezaron a nidificar en ella, aunque en seguida la dejaron. Obviamente, se trataba de una oportunidad aprovechable, aunque al parecer no era especialmente valiosa. En cualquier caso, había una situación problemática. Tal vez otros pájaros resuelvan el problema otro año. En caso contrario, puede que otra caja se muestre más adecuada. Por otro lado, es posible retirar otra caja más adecuada antes de que sea utilizada. El problema de la conveniencia de la caja es claramente objetivo, siendo en parte accidental el que se utilice. Lo mismo ocurre con los nichos ecológicos. Constituyen potencialidades que como tales pueden ser estudiadas de manera objetiva, independientemente hasta cierto punto de la cuestión de si serán actualizadas o no por algún organismo vivo. Un bacteriólogo sabe cómo preparar un nicho ecológico para el cultivo de ciertas bacterias u hongos. Puede ser perfectamente adecuado para ello, aunque el hecho de que sea utilizado y habitado sea otra cuestión.

Una gran parte del tercer mundo objetivo de teorías, libros y argumentos actuales o posibles, surgen como subproducto involuntario de los libros y argumentos realmente producidos. También podemos decir que es un subproducto del lenguaje humano. El propio lenguaje es, como el nido de un pájaro, un subproducto involuntario de acciones orientadas a otros fines.

¿Cómo surge en la jungla una senda de animales? Tal vez algunos animales abran brecha en la maleza para llegar a un abrevadero. Otros animales encuentran más fácil seguir el mismo camino. De este modo, se puede

ampliar y mejorar mediante el uso. No es algo planificado, es una consecuencia involuntaria de la necesidad de moverse con facilidad o ligereza. Es así como se origina una senda —tal vez también las humanas— el lenguaje y otras instituciones útiles, y es así como deben su existencia y desarrollo a su utilidad. No están planificados ni se buscan intencionadamente; tal vez no eran necesarios antes de aparecer. Mas pueden crear una necesidad nueva o un nuevo conjunto de fines: la estructura de orientación a fines de los hombres y animales no está "dada", sino que se desarrolla, con ayuda de algún mecanismo de retroalimentación, más allá de los fines primitivos y de los resultados propuestos o no [5].

De este modo puede surgir todo un nuevo universo de posibilidades o potencialidades —un mundo que en gran medida es autónomo.

Un jardín constituye un ejemplo muy natural. Aún cuando haya podido ser planeado con gran cuidado, por regla general se desarrollará por vías en parte inesperadas. Pero aunque se desarrolle según estaba previsto, algunas interrelaciones inesperadas entre los elementos planificados pueden dar lugar a todo un universo de posibilidades, de nuevos propósitos posibles y de nuevos *problemas*.

El mundo del lenguaje, de las conjeturas, teorías y argumentos —en breve, el universo del conocimiento objetivo— es el más importante de esos universos creados por el hombre que al mismo tiempo son en gran medida autónomos.

La idea de *autonomía* es fundamental para mi teoría del tercer mundo: aunque sea un producto humano, una creación del hombre, a su vez crea, como otros productos animales, su propio *campo de autonomía*.

Hay incontables ejemplos. En la teoría de los números naturales se pueden encontrar los ejemplos tal vez más chocantes que, en todo caso, son los que hay que tener presentes como ejemplos típicos.

Dejando en paz a Kronecker, estoy de acuerdo con Brouwer en que la sucesión de los números naturales es una construcción humana. Pero aunque creemos la sucesión, ésta crea a su vez sus propios problemas autónomos. La distinción entre números pares e impares no es creación nuestra: es una consecuencia involuntaria e inevitable de nuestra creación. Naturalmente, los números primos son hechos autónomos y objetivos igualmente involuntarios; por lo que a ellos se refiere, es obvio que ahí tenemos muchos hechos que *descubrir*: hay conjeturas como la de Goldbach. Además, esas conjeturas, aunque se refieran indirectamente a objetos creados por nosotros, se refieren directamente a problemas y hechos que en cierto modo han surgido a partir de nuestra creación y que no podemos controlar o influenciar: son hechos sólidos y la verdad acerca de ellos es a menudo difícil de descubrir.

Esto ilustra lo que quiero decir cuando afirmo que el tercer mundo es en gran medida autónomo, aunque sea una creación nuestra.

---

[5] Véase Hayek, 1967, cap. 6, especialmente las págs. 96, 100, núm. 12; Descartes, 1637; cf. 1931, pág. 89 [trad. cit., pág. 143]; Popper 1960, pág. 65 [pág. 66-7]; 1966 sección XXIV (i.e., págs. 233-5 más adelante).

Pero la autonomía sólo es parcial: los nuevos problemas llevan a nuevas creaciones o construcciones —como las funciones recursivas o las secuencias de elección libre de Brouwer— que de este modo pueden añadir nuevos objetos al tercer mundo. Además, cada uno de esos pasos creará *nuevos hechos inintencionados, nuevos problemas inesperados* y, a menudo, también *nuevas refutaciones* [*].

Hay también un efecto importantísimo de retroalimentación entre nuestras creaciones y nosotros mismos, entre el tercer mundo y el segundo, ya que los problemas nuevos nos impulsan a crear nuevas cosas.

Se puede describir dicho proceso mediante el siguiente esquema, excesivamente simplificado en cierto modo (véase mi 1966, especialmente pág. 225, más adelante):

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2.$$

Es decir, partimos de un problema P_1, procedemos a una solución tentativa o a una teoría TT que puede resultar equivocada (total o parcialmente); en cualquier caso, estará sujeta a eliminación de errores, EE, que puede consistir en una discusión crítica o en contrastaciones experimentales; sea como sea, de nuestra propia actividad creadora surgen problemas, P_2, que en general no creamos voluntariamente, sino que emergen del campo de nuevas relaciones que no podemos dejar de engendrar con cada acción, por poco que lo pretendamos.

La autonomía del tercer mundo y su retroalimentación sobre el segundo, e incluso sobre el primero, se cuentan entre los hechos más importantes del aumento del conocimiento.

Siguiendo con nuestras consideraciones biológicas, es fácil ver que poseen una importancia general para la teoría de la evolución de Darwin: explican cómo podemos salir de nuestros propios cepos o, para decirlo con una terminología más exquisita, nos ayudan a explicar la "emergencia".

4. Lenguaje, crítica y el tercer mundo

Las creaciones humanas y los efectos de retroalimentación sobre nosotros mismos, especialmente sobre nuestros cerebros, más importantes son las funciones más elevadas del lenguaje humano; concretamente, la *función descriptiva* y la *función argumentadora*.

Los lenguajes humanos comparten con los lenguajes animales las dos funciones lingüísticas inferiores: (1) la autoexpresión y (2) la señalización. La función autoexpresiva o sintomática del lenguaje es obvia: todo lenguaje animal es síntoma del estado de un organismo. La función señalizadora o desencadenadora es igualmente obvia: no decimos que un síntoma es lin-

---

[*] Un ejemplo de esto último es la "refutación por extensión de conceptos" de Lakatos; véase Lakatos 1963-4.

güístico a menos que supongamos que puede desencadenar una respuesta en otro organismo.

Todos los lenguajes animales y todos los fenómenos lingüísticos comparten estas dos funciones inferiores. Pero el lenguaje humano posee además muchas otras funciones [7]. Es muy extraño que casi todos los filósofos hayan pasado por alto alguna de las funciones superiores más importantes. La explicación hay que buscarla en el hecho de que las dos funciones inferiores siempre están presentes cuando tienen lugar las superiores, de modo que siempre es posible "explicar" los fenómenos lingüísticos —en términos de funciones inferiores— como "*expresión*" o "*comunicación*".

Las dos funciones superiores del lenguaje humano más importantes son (3) la función *descriptiva* y (4) la función *argumentadora* [8].

Con la función descriptiva del lenguaje humano emerge la idea reguladora de *verdad*, es decir, la idea de una descripción que encaje con los hechos [9].

Otras ideas reguladoras o evaluativas son las de contenido de verdad y verosimilitud [10].

La función argumentadora del lenguaje humano presupone la función descriptiva: los argumentos versan fundamentamente sobre descripciones desde el punto de vista de las ideas reguladoras de verdad, contenido y verosimilitud.

Pero aquí hay dos cuestiones importantísimas:

(1) Sin el desarrollo de un lenguaje descriptivo exosomático —un lenguaje que, como las herramientas, se desarrolle fuera del cuerpo— nuestra discusión crítica carece de *objeto*. Ahora bien, con el desarrollo de un lenguaje descriptivo (y además, escrito) puede emerger un tercer mundo lingüístico. Los problemas y normas de crítica racional sólo se pueden desarrollar de este modo y sólo en este tercer mundo.

(2) Nuestra humanidad, nuestra razón, se la debemos a este desarrollo de las funciones superiores del lenguaje, ya que nuestros poderes de razonamiento no son más que poderes de argumentación crítica.

Este segundo punto muestra la futilidad de toda teoría sobre el lenguaje

---

[7] Por ejemplo, avisar, exhortar, fingir, etc.

[8] Véase Popper 1963, en especial los capítulos 4 y 12 y las referencias de las páginas 134, 293 y 295 [págs. 158, 339, 341] al libro de Bühler 1934. Bühler fue el primero en discutir la diferencia decisiva que media entre las funciones inferiores y la función descriptiva. Más tarde, merced a mi teoría de la crítica, descubrí la diferencia decisiva que media entre la función descriptiva y la argumentadora. Véase también Popper, 1966, sección XIV y nota 47; véase también la pág. 218 más adelante.

[9] Uno de los mayores descubrimientos de la lógica moderna fue el restablecimiento, debido a Tarski, de la teoría de la verdad como correspondencia (objetiva): verdad=correspondencia con los hechos. Este ensayo se lo debe todo a dicha teoría, aunque, como es natural, no deseo comprometer a Tarski con ninguno de los crímenes aquí perpetrados.

[10] Véase la nota anterior y Popper, 1962a, en especial la pág. 292; también, Popper, 1963, cap. 10 y Apéndice, así como las págs. 51-65 *Supra*, y el capítulo 9, más adelante.

humano que se centre sobre la *expresión y comunicación*. Como veremos, el organismo humano que, como muchas veces se dice, tiende a expresarse a sí mismo, depende en gran medida, por lo que respecta a su estructura, de la emergencia de las dos funciones superiores del lenguaje.

En la evolución de la función argumentadora del lenguaje, la crítica se convierte en el instrumento fundamental del desarrollo ulterior. (La lógica se puede considerar *el organon de la crítica;* véase mi 1963, pág. 64 [Traducción castellana, pág. 79].) El mundo autónomo de las funciones superiores del lenguaje se convierte en el mundo de la ciencia, con lo que el esquema originalmente válido tanto para el mundo animal como para el hombre primitivo,

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$$

se transforma en el esquema del aumento de conocimiento mediante la supresión de errores en virtud de la *crítica racional sistemática*. Se transforma en el esquema de la búsqueda de la verdad y contenido mediante la discusión racional. Describe el modo en que salimos de nuestros cepos. Suministra una descripción racional de la emergencia evolutiva y de nuestra *auto-trascendencia por medio de selección y crítica racional.*

Resumiendo, aunque el significado de "conocimiento" carezca de importancia, como el de cualquier palabra, es importante distinguir sus distintos sentidos.

(1) Conocimiento subjetivo que consta de disposiciones a la acción de carácter innato, junto con sus modificaciones adquiridas.

(2) Conocimiento objetivo, como por ejemplo, el conocimiento científico, que está formado por teorías conjeturales, problemas pendientes, situaciones problemáticas y argumentos.

Todo trabajo científico está dirigido a acrecentar el conocimiento objetivo. Somos trabajadores que colaboramos al aumento del conocimiento objetivo como albañiles que trabajan en una catedral.

Nuestro trabajo, como toda obra humana, es falible. Cometemos errores constantemente, ya que si bien hay criterios objetivos —de verdad, contenido, validez, etc.— podemos transgredirlos.

El lenguaje, la formulación de los problemas, la emergencia de nuevas situaciones problemáticas, las teorías rivales, la crítica mutua mediante argumentos, todos ellos constituyen los medios indispensables del desarrollo científico. Las funciones o dimensiones más importantes del lenguaje humano (que no poseen los lenguajes animales) son las funciones descriptivas y argumentadoras. Aunque sean consecuencias involuntarias de nuestras acciones, como es natural, el desarrollo de estas funciones es cosa nuestra. La argumentación crítica y el conocimiento en sentido objetivo sólo se hacen posibles dentro de un lenguaje enriquecido de este modo.

Difícilmente se pueden exagerar las repercursiones o los efectos de

retroalimentación que produce la evolución del tercer mundo sobre nosotros —nuestros cerebros, tradiciones (si alguien partiese de donde partió Adan no progresaría más que éste), nuestras disposiciones a la acción (es decir, nuestras creencias) [11] y nuestras acciones.

Por el contrario, la *epistemología tradicional* se interesa por el segundo mundo por el conocimiento como un tipo particular de creencia —creencia justificable, como lo es la creencia basada en la percepción—. El resultado es que este tipo de filosofía de la creencia no puede expresar (ni lo intenta siquiera) el fenómeno decisivo de que los filósofos critican sus teorías destruyéndolas por este medio. *Los científicos intentan eliminar sus teorías falsas, intentan dejarlas morir en su rincón. El creyente —hombre o animal— perece junto con sus creencias falsas.*

5. Consideraciones históricas

5.1. *Platón y el neoplatonismo*

Como es bien sabido, Platón fue el descubridor del tercer mundo. Como señalaba Whitehead, toda la filosofía occidental no es más que un conjunto de anotaciones a Platón.

Sólo haré tres consideraciones breves en torno a Platón, dos de las cuales son críticas.

(1) Platón no sólo descubrió el tercer mundo, sino además parte de la influencia o retroalimentación entre el tercer mundo y nosotros: descubrió, no sólo que intentamos captar las ideas de este tercer mundo, sino además que las utilizamos como explicaciones.

(2) El tercer mundo platónico tenía un carácter divino; era inalterable y, naturalmente, verdadero. Por tanto, hay una gran brecha entre éste y mi tercer mundo que es un producto humano cambiante. No sólo contiene teorías verdaderas, sino también falsas y, en especial, problemas pendientes, conjeturas y refutaciones.

Considero que los argumentos (para no hablar de los problemas pendientes) son algunos de los inquilinos más importantes del tercer mundo, aunque Platón, el gran maestro del razonamiento dialéctico, no vio en ellos más que una vía de acceso al tercer mundo.

(3) Platón creía que el tercer mundo de las Formas o Ideas nos suministraba explicaciones últimas (es decir, explicaciones por medio de esencias; véase mi 1963, cap. 3). Así por ejemplo, escribe: "Considero que si nada es bello excepto la idea de belleza absoluta, entonces algo será bello por *la única razón* de que participa de la idea de belleza absoluta. *Este tipo de explicación se aplica a todo*". (Platón, *Fedón,* 100 c.)

---

[11] La teoría de que las creencias podrían calibrarse mediante el grado en que uno está dispuesto a apostar era de sobra conocida en 1771; véase Kant 1778, página 852 [cf. la traducción de Perojo, 1961, vol. II, pág. 194].

Se trata de la teoría de las *explicaciones últimas;* es decir, de explicaciones cuyo explicans ni precisa ni es susceptible de ulteriores explicaciones. Se trata de una teoría de *explicaciones por esencias;* es decir, mediante palabras hipostasiadas.

El resultado fue que Plantón consideró los objetos del tercer mundo algo así como cosas inmateriales o, tal vez, como estrellas o constelaciones —algo que se puede captar e intuir, aunque no es posible tocarlas con la mente. Por eso los inquilinos del tercer mundo —las formas o ideas— se transformaron en conceptos de cosas, esencias o naturalezas de las cosas, más bien que en teorías, argumentos o problemas.

De ahí surgieron consecuencias importantísimas para la historia de la filosofía. De Platón a nuestros días, la mayoría de los filósofos o han sido nominalistas [1 2] o han sido lo que he llamado esencialistas. Se interesan más por el significado (esencial) de las palabras que por la verdad o falsedad de las teorías.

Acostumbro a presentar el problema mediante el siguiente cuadro:

	LAS IDEAS	
	es decir, las	
DESIGNACIONES, TÉRMINOS *o* CONCEPTOS		PROPOSICIONES, ENUNCIADOS *o* TEORÍAS
	pueden formularse con	
PALABRAS		AFIRMACIONES
	que pueden ser	
SIGNIFICATIVAS		VERDADERAS
	y su	
SIGNIFICADO		VERDAD
	se puede reducir a	
CONCEPTOS INDEFINIDOS		PROPOSICIONES PRIMITIVAS
	mediante	
DEFINICIONES		DERIVACIONES
El intento de establecer (más bien que reducir) por estos medios su SIGNIFICADO		VERDAD
lleva a una regresión infinita		

Desde mi punto de vista, *la parte izquierda del cuadro carece de importancia* si se compara con la derecha: lo que debe interesarnos son las teorías, la verdad, los argumentos. Si hay tantos filósofos y científicos que aún

[1 2] Watkins, 1965, cap. VIII, en especial págs. 145 y sigs., y Popper, 1959, páginas 420-2 [págs. 392-5]; 1963, págs. 18 y sigs., 262, 297 y sigs. [págs. 26 y sigs., 303 y 344].

consideran que los conceptos y sistemas conceptuales (junto con los problemas de su significado o el significado de las palabras) tienen una importancia comparable a la de las teorías y sistemas teóricos (junto con el problema de su verdad o la verdad de los enunciados), es que siguen presa del error fundamental de Platón [13], puesto que los conceptos son en parte medios para formular teorías y, en parte, medios para resumir teorías. En cualquier caso, su significación tiene un carácter fundamentalmente instrumental y siempre pueden ser sustituidos por otros conceptos.

Los contenidos y objetos de pensamiento parecen haber desempeñado un papel importante en el Estoicismo y Neo-platonismo: Plotino mantuvo la división platónica entre el mundo empírico y el mundo platónico de las Formas o Ideas. Sin embargo Plotino, como Aristóteles [14], destruyó la trascendencia del mundo platónico situándolo en la conciencia divina.

Plotino criticó a Aristóteles por no ser capaz de distinguir entre la Primera Hipóstasis (Unidad) y la Segunda Hipóstasis (el entendimiento divino). Sin embargo, siguió a Aristóteles al identificar los actos de pensamiento de Dios con sus propios contenidos u objetos y elaboró además este punto de vista, considerando las Formas o Ideas del mundo inteligible platónico como estados inmanentes de conciencia del entendimiento divino [15].

5. 2. *Hegel*

Hegel era un tanto platónico (o neo-platónico) y, como el mismo Platón, un tanto heraclíteo. Era un platónico cuyo mundo de las ideas cambiaba y evolucionaba. Las "Formas" o "Ideas" platónicas eran objetivas, sin que tuviesen nada que ver con las ideas conscientes de una mente subjetiva; habitaban un mundo celeste (supralunar en sentido aristotélico) divino y estable. Por el contrario, las ideas de Hegel, como las de Plotino, eran fenómenos conscientes: pensamientos que se pensaban a sí mismos, alojados en un cierto tipo de conciencia, mente o "Espíritu", junto con el cual cambiaban

---

[13] El error, ya tradicional, se conoce como "el problema de los universales". Se podría sustituir por "el problema de las teorías" o "el problema del contenido teórico de todo lenguaje humano". Véase Popper 1959, secciones 4 (con la nueva nota * 1) y 25.

Dicho sea de paso, es evidente que de las tres famosas posiciones —*universale ante rem, in re* y *post rem*— la última, con su significado habitual, va en contra del tercer mundo al tratar de explicar el lenguaje como expresión, mientras que la primera (platónica) es favorable al tercer mundo. Es bastante interesante el que la posición (aristotélica) intermedia (*in re*) pueda o bien ser contraria al tercer mundo o bien ignorar el problema. De este modo pone de manifiesto la confusa influencia del conceptualismo.

[14] Cf. Aristóteles, *Metafísica*, XII (A) 7: 1072b 21 y sigs.; y 9: de 1974b 15 a 1075a 4. Este pasaje (resumido por Ross con las palabras: "el pensamiento divino debe ocuparse del objeto divino por excelencia que es él mismo") contiene una crítica implícita a Platón. Su afinidad con las ideas platónicas se ve con especial claridad en las líneas 25 y sigs.: "Piensa en lo más divino y preciado que no cambia, ya que el cambio habría de ser para algo peor..." (Véase también, Aristóteles, *De Anima*, 429b 27 y siguientes, en especial 430a 4.)

[15] Cf. Plotino, *Eneadas*, II. 4.4 (1883, pág. 153, 3); III. 8.11 (1883, pág. 346, 6); V. 3.2-5; V. 9.5-8; VI. 5.2; VI. 6.6-7.

y evolucionaban. El hecho de que el "Espíritu Objetivo" y el "Espíritu Absoluto" hegelianos estuviesen sometidos a cambios constituye el único aspecto en que sus Espíritus se asemejan más a mi tercer mundo que el mundo de las Ideas platónicas (o el mundo de los "enunciados en sí mismos" de Bolzano).

Las diferencias más importantes existentes entre el "Espíritu Objetivo" o el "Espíritu Absoluto", por un lado, y mi "tercer mundo", por el otro, son las siguientes:

(1) Según Hegel, el hombre no es creador, a pesar de que el "Espíritu Objetivo" (que engloba la creación artística) y el Absoluto (que engloba la filosofía) están formados por productos humanos. Es el Espíritu Objetivo hipostasiado, es la divina conciencia de sí del Universo lo que mueve al hombre: "los individuos... son instrumentos", instrumentos del espíritu de la época y su trabajo, su "quehacer substancial", es "dispuesto y fijado independientemente de ellos". (Cfr. Hegel 1830, parágrafo 551). Asi pues, lo que he denominado la autonomía del tercer mundo y su efecto de retroalimentación se hace omnipotente con Hegel: se trata tan sólo de uno de los aspectos de su sistema en que se pone al descubierto su trasfondo teológico. Frente a ello, sostengo que el elemento creador individual, la relación de intercambio entre el hombre y su obra es de la mayor importancia. En el caso de Hegel esta situación degenera en la doctrina de que los grandes hombres son una especie de mediums en los que se expresa el espíritu de la época.

(2) A pesar de cierta semejanza superficial, entre la dialéctica hegeliana y mi esquema evolucionista

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$$

hay una diferencia fundamental. Mi esquema funciona con la supresión de errores y, a nivel científico, con la crítica consciente presidida por la idea reguladora de la búsqueda de la verdad.

Evidentemente, la crítica consiste en buscar y eliminar contradicciones: las dificultades creadas por la exigencia de suprimirlas constituyen el nuevo problema (P_2). De este modo, la supresión del error conduce al crecimiento objetivo de nuestro conocimiento —en sentido objetivo—. Conduce al aumento de la verosimilitud objetiva: permite la aproximación a la verdad (absoluta).

Por otra parte, Hegel es un relativista [16]. No considera que nuestra misión sea buscar contradicciones para suprimirlas, pues cree que las contradiciones son tan excelentes como los sistemas teóricos no-contradictorios (si no mejores) por la razón de que suministran el mecanismo mediante el que el Espíritu se impulsa a sí mismo. Así pues, la crítica racional no des-

---

[16] Véase Popper, 1963, cap. 15; Popper, 1962, Apéndice al vol. II: "Facts, Standars and Truth: A further Criticism of Relativism". [Dicho apéndice de la tercera edición no aparece en la traducción castellana hecha sobre la segunda edición inglesa.]

empeña ningún papel en el automatismo hegeliano, cosa que también ocurría con la creatividad humana [17].

(3) Mientras que Platón hacía que sus ideas hipotasiadas habitasen en un mundo celeste, Hegel personaliza su Espíritu en una especie de conciencia divina: Las Ideas habitan en ella del mismo modo que las ideas humanas habitan en una conciencia humana. Por todo esto, su doctrina afirma que el Espíritu, no sólo es consciente, sino que también es un sujeto. Por el contrario, mi tercer mundo no posee ninguna semejanza con la conciencia humana; aunque sus primeros inquilinos sean productos de la conciencia humana, son totalmente distintos de las ideas conscientes o pensamientos en sentido subjetivo.

5.3. *Bolzano y Frege*

Los enunciados en sí mismos y las verdades en sí mismas de Bolzano son, evidentemente, inquilinos de mi tercer mundo. Pero no estaba nada claro para él cuáles eran sus relaciones con el resto del mundo [18].

En cierta medida, he intentado resolver la dificultad central de Bolzano comparando la condición y autonomía del tercer mundo con los productos de los animales y señalando también su origen en las funciones superiores del lenguaje humano.

Por lo que a Frege respecta, no hay ninguna duda acerca de su clara distinción entre los actos de pensamiento subjetivos, o pensamientos en sentido subjetivo, y el pensamiento objetivo o contenido del pensamiento [19].

Hay que admitir que su interés por las cláusulas subordinadas de las oraciones y por el estilo indirecto le ha hecho el padre de la lógica epistémica moderna [20]. Pero no creo que le afecten las críticas a la lógica epistémica que haré más adelante (véase más abajo, la sección 7): que yo sepa, no pensaba en estos contextos epistemológicos en el sentido de una teoría del conocimiento científico.

5.4. *Empirismo*

El empirismo —digamos el de Locke, Berkeley y Hume— ha de entenderse en su contexto histórico: su problema fundamental era, sencillamente, religión *versus* irreligión; o, más exactamente, la justificación racional —o justificabilidad— del Cristianismo comparado con el conocimiento científico.

Esto explica por qué consideraban el conocimiento como un tipo de

---

[17] Véase Lakatos, 1963-4, pág. 234, nota 1 (separata, pág. 59).

[18] Bolzano 1837, vol. I, 19, pág. 78, dice que los enunciados (y verdades) en sí mismos carecen de ser ("*Dasein*"), existencia o realidad. Sin embargo, añade que un enunciado en sí mismo *no* es *simplemente* "algo enunciado que presuponga, por tanto, una persona que lo haya enunciado".

[19] Véase la cita que aparece en la pasada sección 1 de Frege, 1892, pág. 32 [página 58, nota 5], y Frege, 1894.

[20] El camino va de Frege a Russell, 1922, pág. 19, y Wittgenstein, 1922, 5.542.

creencia —creencia justificada por elementos de juicio, especialmente perceptivos o de los sentidos.

Aunque sus posturas relativas a las relaciones entre ciencia y religión difieren ampliamente, Locke, Berkeley [21] y Hume están esencialmente de acuerdo en la exigencia (que Hume considera a veces como un ideal inalcanzable) de rechazar cualquier proposición —especialmente las que tienen un alcance existencial— que carezca de elementos de juicio suficientes en su apoyo, debiéndose aceptar solamente aquellas que gocen de elementos de juicio sufientes en su favor: que puedan ser probadas o verificadas por la evidencia de los sentidos.

Esta posición se puede analizar de diversas maneras. Un análisis un tanto superficial vendría dado por la siguiente cadena de ecuaciones o equivalencias, la mayoría de las cuales pueden apoyarse en pasajes de empiristas ingleses, incluso de Bertrand Russell [22].

p es verificado o demostrado por la experiencia de los sentidos = estamos autorizados a, o poseemos razones suficientes para, creer p = creemos, juzgamos, afirmamos, asentimos o sabemos que p es verdadero = p es verdadero = p.

Lo notable de esta postura, que *confunde los elementos de juicio o demostración con el aserto que hay que probar,* es que quien la mantiene debe *rechazar la ley del tercio excluso.* Es obvio que puede surgir un caso (de hecho, esto será lo que ocurra normalmente), en que los elementos de juicio válidos no apoyen o demuestren plenamente ni p ni no-p. Sin embargo, parece que nadie se dió cuenta de esto hasta Brouwer.

La omisión del rechazo de la ley del tercio excluso es particularmente sorprendente en el caso de Berkeley, ya que si

$$esse = percipi,$$

entonces la verdad de un enunciado acerca de la realidad sólo se puede establecer mediante enunciados perceptivos. Sin embargo, Berkeley sugiere en sus Diálogos [23] —de un modo muy semejante a como hacía Descartes—

---

[21] Para la postura de Berkeley, confróntese Popper, 1963, sección 1 del capítulo 3, así como el capítulo 6.

[22] Cf. Russell, 1906-7, pág. 54: "La verdad es una cualidad de las creencias"; Russell, 1910: "Emplearé las palabras 'creencia' y 'juicio' como sinónimas" (página 172, nota [trad., pág. 215 nota 1].); o "...el juicio es... una relación múltiple del espíritu con los otros varios términos de que se ocupa el juicio" (pág. 180 [pág. 225]). También sostiene que "la percepción siempre es verdadera (incluso en sueños y alucinaciones)" (pág. 181 [pág. 227]. O cf. Russell 1959, pág. 183:... "mas desde el punto de vista de la teoría del conocimiento y de la definición de verdad, lo que importa son las oraciones que expresan creencias". Véase también Russell, 1922, páginas 19 y sigts. [págs. 17 y sigts.], y las *"aptitudes epistémicas"* de Ducasse, 1940, páginas 701-11. Está claro que tanto Russell como Ducasse forman parte de este tipo de epistemólogos tradicionales que estudian el conocimiento en sentido subjetivo o del segundo mundo. Esta tradición desborda cumplidamente el empirismo.

[23] Véase el segundo diálogo entre Hylas y Philonous (Berkeley, 1949, pág. 218, líneas 15 y sigs. [Véase la traducción de Viqueira, pág. 94]): "Para mí es razón suficiente para no creer en la existencia de algo el que yo vea que no hay razón para creerlo". Cf. Descartes, 1637, parte IV (primer párrafo) [cf. la trad. citada, pág. 156.]:

que hemos de rechazar *p* si "no hay razones para creerlo". La ausencia de tales razones puede ser compatible, no obstante, con la creencia de razones para creer no-*p*.

6. Evaluación y crítica de la epistemología de Brouwer

En esta sección es mi deseo rendir homenaje a L. E. J. Brouwer [24].

Sería una presunción intolerable por mi parte intentar elogiar, y más aún criticar, los trabajos matemáticos de Brouwer. Con todo, se me permitirá intentar criticar su epistemología y su filosofía de las matemáticas intuicionista. Si me atrevo a ello, es sólo con el deseo de contribuir, por poco que sea, a la clarificación y ulterior desarrollo de las ideas de Brouwer.

En su Lección Inaugural (1912) Brouwer parte de Kant. Dice que la filosofía intuicionista kantiana de la geometría —su doctrina de la intuición pura del espacio— ha de ser abandonada a la luz de la geometría no-euclídea. Pero eso, añade Brouwer, no es preciso puesto que podemos aritmetizar la geometría: podemos adoptar honradamente nuestra posición, por lo que respecta a la tesis kantiana de la aritmética y a su doctrina de que la aritmética se basa en la intuición pura del tiempo.

Me da la impresión de que ya no se puede sostener esta posición de Brouwer, puesto que si decimos que la geometría no-ouclídea destruye la teoría kantiana del espacio, nos vemos obligados a afirmar que la relatividad especial destruye su teoría del tiempo. Esto es así, porque Kant dice explícitamente que sólo hay *un* tiempo, siendo absolutamente decisiva la idea intuicionista de simultaneidad (absoluta) [25].

Se podría argüir —siguiendo un camino paralelo al tomado por Heyting [26] en unas consideraciones que hace sobre esta cuestión— que Brouwer no habría desarrollado sus ideas epistemológicas y filosóficas relativas a las matemáticas intuicionistas si hubiese tenido conocimiento en su día de la analogía entre la relativización del tiempo de Einstein y la geometría no-euclídea. Parafraseando a Heyting, hubiese sido una desgracia.

"Era necesario... que rechazase como completamente falsa ["*aperte falsa*" en la versión latina] cualquier opinión en la que pudiese encontrar la menor razón para la duda".

[24] Esta sección sobre Brouwer la hice para rendir homenaje a este gran matemático y filósofo, muerto poco antes de que comenzase el Congreso en el que iba a leer este trabajo. A quienes no estén familiarizados con la filosofía intuicionista de las matemáticas de Brouwer (y Kant), les resultará más cómodo saltar esta sección para continuar con la 7, más adelante.

[25] En la Estética Trascendental (Kant, 1778, págs. 46 y sigs.; traducción de Kemp-Smith, págs. 74 y sigs. [traducción de Perojo, vol. I, págs. 183 y sigs.]) Kant subraya en el punto 1 el carácter *a priori* de la simultaneidad; en los puntos 3 y 4, que sólo puede haber *un* tiempo y, en el 4, que el tiempo *no es un concepto discursivo*, sino una forma pura de la intuición" (o, más exactamente, *la* forma pura de la intuición sensible). En el último párrafo, antes de la Conclusión, en la pág. 72 (Kemp-Smith, pág. 90 [trad cit., pág. 200]) dice explícitamente que las intuiciones de espacio y tiempo no son intelectuales.

[26] Véase la cita de Heyting que aparece, más arriba, en la sección 1.

Con todo, no es probable que Brouwer se hubiese sentido muy impresionado por la relatividad especial. Podría haber dejado de citar a Kant como precursor de su intuicionismo, conservando su propia teoría de un tiempo *personal*; un tiempo de nuestra propia experiencia íntima e inmediata (véase Brouwer, 1949) que no se vería afectada por la relatividad como se vio la teoría de Kant.

Así pues, no es necesario considerar kantiano a Brouwer. Sin embargo, no es fácil desligarlo totalmente de Kant, porque la idea que tiene Brouwer de la intuición y su uso del término "intuición" no se puede comprender plenamente sin analizar su trasfondo kantiano.

Para Kant, *la intuición es una fuente de conocimiento;* además, la intuición "pura" ("la intuición pura de espacio y tiempo") constituye una fuente de conocimiento infalible de la que surge la *certeza absoluta*. Esto es de la mayor importancia para comprender a Brouwer que adopta con toda claridad esta doctrina epistemológica de Kant.

La doctrina en cuestión tiene su historia. Kant la tomó de Plotino, Santo Tomás, Descartes y otros. Como es evidente, en su origen, intuición significaba percepción: es lo que vemos o percibimos si miramos o prestamos atención a un objeto. Mas a partir por lo menos de Plotino, se desarrolla la separación entre *intuición* por una parte y pensamiento *discursivo* por otra. La intuición es el modo que tiene Dios de conocerlo todo de una ojeada, en un instante, atemporalmente. El pensamiento discursivo es el modo de conocer humano: como en un discurso, razonamos por pasos, lo cual lleva tiempo.

Ahora bien, Kant sostuvo (contra Descartes) la doctrina de que no poseemos una facultad de intuición intelectual, razón por la cual nuestro entendimiento —nuestros conceptos— son vacíos o analíticos si no se aplican a una materia que, o bien nos suministran nuestros sentidos (intuición sensible), o bien viene dada por "*conceptos construidos en la intuición pura de espacio y tiempo*" [27]. Sólo de este modo podemos obtener conocimiento sintético *a priori*: nuestro entendimiento es esencialmente discursivo; se ve obligado a proceder con lógica que es vacía— "analítica".

Según Kant, la intuición sensible presupone la intuición pura: nuestros sentidos no pueden funcionar sin ordenar las percepciones en el marco suministrado por el espacio y el tiempo. Así pues, el espacio y el tiempo son previos a toda intuición sensible, siendo válidas *a priori* las teorías del espacio y el tiempo —geometría y aritmética—. La fuente de su validez *a priori* es la facultad humana de una *intuición pura* que se limita estricta-

[27] Véase Kant, 1778, pág. 741 [trad. Perojo-Klein, vol. III, pág. 324]: "construir un concepto significa exponer *a priori* la intuición [la "intuición pura"] que le corresponde". Véase también la pág. 747 [*ibid.*, pág. 327]: "En el ejemplo señalado, nos hemos limitado a poner de manifiesto la gran diferencia que hay entre el uso discursivo de la razón, al proceder por conceptos, y el uso intuitivo fundado en la construcción de los objetos". En la pág. 751 [pág. 330] se explica más ampliamente la "*construcción de conceptos*": podemos determinar *a priori* nuestros conceptos en la intuición, creando los *objetos mismos* por síntesis uniforme en el espacio y en el tiempo". (El subrayado es mío en parte.)

mente a este campo y que se distingue tajantemente del modo de pensar intelectual o discursivo.

Kant sostuvo la teoría de que los *axiomas de las matemáticas* se basan en la intuición pura (Kant, 1778, págs. 760 y sigs. [Trad. Losada cit. páginas 335 y sigs.]): su verdad se puede "ver" o "percibir" de un modo no-sensible. Además, la intuición pura estaba presente en *todos los pasos de todas las demostraciones geométricas* (y matemáticas, en general) [28]: para seguir una demostración es preciso mirar una figura (dibujada). Este "mirar" no pertenece a la intuición sensible, sino a la pura, como demuestra el hecho de que muchas veces la figura puede ser convincente aunque esté dibujada de un modo muy burdo y como demuestra el hecho de que un triángulo pueda representarnos, con *un* sólo dibujo, una infinitud de variantes posibles —triángulos de todas las formas y tamaños.

Consideraciones análogas se aplican a la aritmética que, a su vez, se basa esencialmente en la intuición pura del tiempo.

Ahora bien, esta teoría acerca de las fuentes del conocimiento matemático se ve afectada, en su versión kantiana, por una dificultad grave. Aunque admitimos todo lo que dice Kant, nos quedamos perplejos, ya que es obvio que la geometría euclídea, utilice o no la intuición pura, recurre a argumentos intelectuales, a la deducción lógica. *No se puede negar que la matemática emplea el pensamiento discursivo.* El discurso de Euclides avanza paso a paso a través de proposiciones y de libros enteros: no fue concebido en un único instante intuitivo. Aún cuando admitamos, en aras del argumento, el carácter indispensable de la intuición pura *en cada uno de los pasos individuales sin excepción* (cosa que a los modernos nos resulta difícil de admitir), el carácter progresivo, lógico y discursivo de las derivaciones de Euclides es tan inconfundible y ha sido tan universalmente conocido e imitado (Spinoza, Newton), que resulta difícil creer que Kant pueda haberlo ignorado. De hecho, Kant sabía todo esto probablemente tan bien como cualquiera, pero se vió obligado a adoptar su postura por (1) la estructura de la *Crítica* en la que la "Estética trascendental" precede a la "Lógica Trascendental" y por (2) su separación tajante (yo diría que insostenible-mente tajante) entre pensamiento intuitivo y discursivo. Tal como Kant lo plantea, uno se inclinaría a pensar que no hay simplemente una laguna, sino una contradicción en su exclusión de los argumentos discursivos de la geometría y la aritmética.

Al llenar esa laguna, Brouwer mostró que no era así. Me refiero a su teoría sobre *la relación que hay entre las matemáticas, por un lado, y el lenguaje y la lógica, por otro.*

---

[28] Cf. Kant, 1778, págs. 741-64 [trad. cit., vol. II, págs. 323-38]. Véase, por ejemplo, el final de la página 762 [pág. 336], donde dice que las demostraciones matemáticas ("incluso en álgebra"): "previenen de errores a todos los raciocinios, poniendo a la vista cada uno de ellos". Cf. también, por ejemplo, el comienzo de la página 745 [pág. 326] donde Kant habla de "encadenamiento de conclusiones guiado siempre por la intuición". (En el mismo pasaje —pág. 748 [pág. 328]— "construir" se explica como "representar en la intuición".)

Brouwer resolvió el problema estableciendo una distinción tajante entre las *matemáticas en cuanto tales y su expresión lingüística o comunicación.* Consideraba las matemáticas en sí mismas como una actividad extralingüística consistente esencialmente en una construcción mental sobre la base de la intuición pura del tiempo. Por medio de esta construcción, creamos en nuestra intuición, en nuestra mente, los objetos de la matemática que posteriormente —tras su creación— podemos intentar describir y comunicar a los demás. Por tanto, la descripción lingüística y el argumento discursivo con su lógica vienen después de la actividad esencialmente matemática: vienen siempre precedidos por la construcción de un objeto matemático —como por ejemplo, una demostración.

Esto resuelve el problema que habíamos descubierto en la *Crítica* de Kant. Lo que a primera vista parece una contradicción kantiana se elimina de un modo muy ingenioso mediante la doctrina de la separación tajante de dos niveles, uno intuitivo y mental, esencial para el pensamiento matemático, y otro discursivo y lingüístico, esencial tan sólo para la comunicación.

Como toda gran teoría, la de Brouwer demuestra su valor por su fertilidad. Resolvió de un plumazo tres grandes conjuntos de problemas de la filosofía de las matemáticas:

(1) *Problemas epistemológicos* relativos a la fuente de la certeza matemática, la naturaleza de la evidencia matemática y la naturaleza de las demostraciones matemáticas. La resolución de estos problemas se llevó a a cabo, respectivamente, mediante la doctrina de la intuición como fuente de conocimiento, la doctrina según la cual podemos ver intuitivamente los objetos matemáticos que hemos construido y la doctrina según la cual una demostración matemática es una construcción secuencial o una construcción de construcciones.

(2) *Problemas ontológicos* relativos a la naturaleza de los objetos matemáticos y a la naturaleza de su modo de existir. Estos problemas fueron resueltos con una doctrina que tiene dos vertientes: por un lado, era *constructivista* y por el otro, *mentalista,* por lo cual situaba todos los objetos matemáticos en lo que yo llamo "segundo mundo". Los objetos matemáticos eran construcciones de la mente humana y existían sólo como construcciones situadas en la misma. Su objetividad —su carácter de objetos y la objetividad de su existencia— descansaba únicamente en la posibilidad de repetir su construcción a voluntad.

Así, podía Brouwer dar a entender en su Lección Inaugural que, para el intuicionista, los objetos matemáticos existen en la mente humana, mientras que para el formalista existen "sobre el papel" [29].

---

[29] Cf. el final del tercer párrafo de Brouwer, 1912. Allí, Brouwer habla de la existencia, no de la matemática, sino de la "exactitud matemática" y, *tal como está planteado,* el pasaje se aplica consecuentemente a los problemas (1) y (3) de un modo aún más directo que al problema ontológico (2). Pero, no cabe duda de que también pretende aplicarse a (2). Dice el pasaje, en la traducción de Dresden: "El problema de la existencia de la exactitud matemática se resuelve de distinto modo... El intuicionista dice: en el entendimiento humano; el formalista dice: sobre el papel".

(3) *Problemas metodológicos* relativos a las demostraciones matemáticas.

Ingenuamente, podemos distiguir dos modos fundamentales de interesarse por las matemáticas. Un matemático puede interesarse básicamente por los teoremas, por la verdad o falsedad de las proposiciones matemáticas. Otro matemático puede interesarse sobre todo por las pruebas: por el problema de la existencia de demostraciones de uno u otro teorema, así como por el carácter de las pruebas. Si predomina el primer tipo de interés (como parece ocurrir con Polya, por ejemplo), suele darse una conexión con intereses por el descubrimiento de "hechos" matemáticos y, por tanto, con la platonización de la heurística matemática. Si predomina el segundo tipo de intereses, entonces las demostraciones no son simples medios de asegurar los teoremas relativos a los objetos matemáticos, sino que son, en sí mismas, objetos matemáticos. Creo que era éste el caso de Brouwer: las construcciones que constituían pruebas no sólo creaban y establecían objetos matemáticos, sino que, al mismo tiempo, ellas mismas eran objetos matemáticos —tal vez, incluso los más importantes—. Así, afirmar un teorema significa afirmar la existencia de su demostración y negarlo significa afirmar la existencia de una refutación; es decir, una prueba de su carácter absurdo. Esto condujo inmediatamente al rechazo brouweriano de la ley del tercio excluso y de las pruebas indirectas, y a la exigencia de demostrar la existencia sólo mediante una construcción real —haciéndola visible, como si dijéramos— del objeto matemático en cuestión.

También le llevó a Brouwer a rechazar el "platonismo", entendido como la doctrina según la cual los objetos matemáticos poseen lo que yo llamo un modo de existencia "autónomo"; es decir, que puedan existir sin que nosotros los hayamos construido y, por tanto, sin que su existencia haya sido demostrada.

Hasta aquí, he intentado comprender la epistemología de Brouwer, suponiendo básicamente que surge de un intento de resolver una dificultad de la filosofía de las matemáticas de Kant. Procederé ahora a llevar a cabo lo que anunciaba en el título de la sección —a valorar y criticar la epistemología de Brouwer.

Desde el punto de vista de este escrito, uno de los mayores logros de Brouwer es haber visto que las matemáticas —y, tal vez, podría añadir, el tercer mundo— son una creación del hombre.

Esta idea es tan radicalmente anti-platónica que es comprensible el hecho de que Brouwer no se diese cuenta de que se puede combinar con cierto tipo de platonismo. Me refiero a la doctrina de la *autonomía* (parcial) de las matemáticas y del tercer mundo, tal como sugeríamos más arriba en la sección 3.

Otros de los grandes logros de Brouwer, desde un punto de vista filosófico, es su antiformalismo: el reconocimiento de que los objetos matemáticos han de existir antes de que se pueda hablar de ellos.

Pero volvamos a la crítica de la solución de Brouwer a los tres grandes tipos de problemas de la filosofía de las matemáticas ya discutidos en la presente sección.

(1') *Problemas epistemológicos*: Intuición en general y la teoría del tiempo en particular.

No intento cambiar el nombre de "Intuicionismo". Puesto que no hay duda de que el nombre será conservado, es más importante eliminar la equivocada filosofía de la intuición como fuente infalible del conocimiento.

No hay fuentes de conocimiento terminantes y ninguna "fuente" es especialmente fiable [30]. Cualquier cosa es bien recibida como fuente de inspiración, incluso la "intuición"; especialmente si nos sugiere nuevos problemas. Pero nada es seguro y todos somos falibles.

Además, la distinción tajante que hace Kant entre el pensamiento discursivo e intuitivo es insostenible. Sea lo que sea la "intuición", es en gran medida un producto de nuestro desarrollo cultural y de nuestros esfuerzos con el pensamiento discursivo. La idea kantiana de un tipo fijo de intuición pura compartido por todos nosotros (tal vez no por los animales, a pesar de su equipo perceptivo semejante) no se puede aceptar fácilmente, ya que tras un entrenamiento con el pensamiento discursivo, nuestra aprehensión intuitiva se hace claramente distinta de como era antes.

Todo esto se aplica a nuestra intuición del tiempo. Personalmente, encuentro convincente el informe que hace Benjamin Lee Whorf de los indios Hopi[31] y su intuición del tiempo manifiestamente distinta de la nuestra. Mas, aunque su informe fuese incorrecto (cosa que considero improbable) nos muestra posibilidades que ni Kant ni Brouwer se pararon nunca a considerar. Si Whorf estuviese en lo cierto, entonces nuestra aprehensión intuitiva del tiempo —el modo que tenemos de "ver" las relaciones temporales— dependería en parte de nuestro lenguaje y de las teorías y mitos que lleva incorporados: *nuestra propia intuición europea del tiempo tiene muchas deudas contraídas con los orígenes griegos de nuestra civilización, con su énfasis en el pensamiento discursivo.*

En cualquier caso, nuestra intuición del tiempo puede cambiar, si cambian nuestras teorías. Las intuiciones de Newton, Kant y Laplace difieren de las de Einstein, y el papel del tiempo en la física de las partículas difiere del que tiene en la física del continuo, especialmente en óptica. Mientras que la física de las partículas sugiere instantes inextensos, tajantemente separados, un "*punctum temporis*" que separa el pasado del futuro y, por tanto, unas coordenadas temporales compuestas por (un continuo de) instantes inextensos y un mundo cuyos "estados" pueden determinarse en cada uno de esos instantes inextensos, la situación en óptica es muy distinta. Del mismo modo que en óptica hay retículos cuyas partes actúan conjuntamente a lo largo de grandes distancias, también hay eventos temporalmente exten-

[30] He tratado extensamente este problema en mi conferencia "Sobre las fuentes del conocimiento y la ignorancia" que ahora constituye la Introducción a Popper, 1963.

[31] Cf. "Un modelo indio-americano del Universo", en Whorf, 1966.

sos (ondas con su frecuencia) cuyas partes actúan conjuntamente a lo largo de considerables distancias temporales. Así mismo, por lo que respecta a la óptica, *no puede haber en física un estado del mundo en un instante temporal*. Este argumento podría establecer, y establece, una gran diversidad en nuestra intuición: el presente psicológico que ha sido considerado como una apariencia, ni es aparente ni se confina a la psicología, sino que es algo genuino que tiene ya su lugar en física [33].

Así pues, la doctrina de la intuición como fuente infalible de conocimiento, no sólo es un mito, sino que además nuestra intuición del tiempo está especialmente sujeta a crítica y modificación, del mismo modo que ocurre, según el propio Brouwer admite, con nuestra intuición del espacio.

El punto de vista fundamental que aquí expongo se lo debo a la filosofía de las matemáticas de Lakatos, según el cual las matemáticas (y no sólo las ciencias naturales) se desarrollan a través de la crítica de las conjeturas y de audaces pruebas informales. Esto presupone que dichas conjeturas y pruebas se formulan lingüísticamente, situándolas, por tanto, en el tercer mundo. El lenguaje, que en un principio no era más que un medio de comunicar descripciones de objetos preligüísticos, se convierte de este modo en una *parte esencial* de la tarea científica, incluso en las matemáticas que, a su vez, pasan a formar parte del tercer mundo. Además, hay estratos o niveles en el lenguaje (estén o no formalizados en una jerarquía de metalenguajes).

Si la epistemología intuicionista fuese correcta, no sería un problema la competencia matemática. (Si la teoría kantiana fuese correcta, sería ininteligible por qué nosotros —o más exactamente, Platón y su escuela— hemos tenido que esperar tanto tiempo por Euclides [33].) Sin embargo, surge el problema de que incluso matemáticos intuicionistas muy competentes pueden estar en desacuerdo en algunos puntos difíciles [34]. Dado un desacuerdo, no es preciso investigar qué alternativa es la correcta. Basta con señalar que, una vez que se puede criticar una construcción intuicionista, el problema planteado sólo se puede resolver *utilizando esencialmente un lenguaje argumentador*. Evidentemente, el uso crítico esencial del lenguaje no nos compromete a emplear argumentos sancionados por la matemática intuicionista (aunque, como veremos, esto plantea un problema). Por el momento, lo que quiero decir es tan sólo esto: tan pronto como se puede poner en cuestión la admisibilidad de una construcción matemática propuesta por el intuicionismo —como es obvio que se puede— el lenguaje

---

[33] Cf. Gombrich 1964, especialmente la pág. 297: "Si queremos desarrollar esta idea hasta su conclusión lógica, el *punctum temporis* ni siquiera puede aparecer como un punto mínimo, ya que la luz tiene una frecuencia". (El razonamiento puede sostenerse tomando en cuenta condiciones límite.)

[33] Cf. la consideración correspondiente sobre la visión apriorística kantiana acerca de la física de Newton en Popper, 1963, cap. 2, el párrafo al que pertenece la nota 63 [trad. págs. 113-4].

[34] Cf. los comentarios de S. C. Kleene en Kleene y Vesley 1965, págs. 176-83, acerca de Brouwer, 1951, págs. 357-8, a quien Kleene critica a la luz de la nota de Brouwer de la página 1248 de Brouwer, 1949.

pasa a ser algo más que un simple medio de comunicación del que podamos prescindir en principio; es más bien el medio indispensable para la discusión crítica. De acuerdo con ello, ya no es la construcción intuicionista la única "que es objetiva en el sentido de ser irrelevante el sujeto que hace la construcción" [35]; por el contrario, la objetividad, incluso la de la matemática intuicionista, descansa en la criticabilidad de sus argutos, como ocurre con todas las ciencias. Ahora bien, lo que esto significa es que el lenguaje se hace indispensable como medio para argumentar, para discutir críticamente [36].

Por esta razón, considero equivocada la epistemología subjetivista de Brouwer y la justificación filosófica de su matemática intuicionista. Hay un intercambio entre construcción, crítica, "intuición" e incluso tradición que él no es capaz de tener en cuenta.

No obstante, estoy dispuesto a admitir que Brouwer lleva parte de razón en su concepción errónea del papel del lenguaje. Aunque la objetividad de cualquier ciencia, incluso de las matemáticas, está inseparablemente ligada a su criticabilidad y, por tanto, a su formulación lingüística, Brouwer hacía bien en reaccionar violentamente en contra de la tesis según la cual las matemáticas *no son más que* un juego del lenguaje formal o, en otras palabras, que no existen objetos matemáticos extra-lingüísticos; es decir, pensamientos (o contenidos de pensamientos, para decirlo de un modo más exacto desde mi punto de vista). Como él subrayaba, las matemáticas hablan *acerca de* esos objetos y, en este sentido, el lenguaje matemático es secundario respecto a esos objetos. Pero esto no quiere decir que podamos construir las matemáticas sin un lenguaje: no puede haber construcción sin un control crítico constante ni crítica sin una formulación lingüística de nuestras construcciones, tratándolas como objetos del tercer mundo. Aunque el tercer mundo no es idéntico al mundo de las formas lingüísticas, surge junto con el lenguaje argumentador: es un subproducto del lenguaje. Esto explica por qué cuando nuestras construcciones se vuelven problemáticas, también se puede tornar problemático el lenguaje sistematizado y axiomatizado y por qué la formalización se puede convertir en una rama de la construcción matemática. Creo que es esto lo que el profesor Myhill quiere decir cuando afirma que "*nuestras formalizaciones corrigen nuestras intuiciones, a la vez que nuestras intuiciones* conforman nuestras formalizaciones" [37]. Lo que confiere un valor particular a esta consideración es el hecho de que, habiendo sido hecha en relación con la prueba intuicionista brouweriana, parece aportar además una corrección a la epistemología de Brouwer.

(2') *Problemas ontológicos*: El propio Brouwer se dió cuenta algunas veces de que los objetos matemáticos deben en parte su existencia al lenguaje. Así, escribió en 1924: "Las Matemáticas se basan en ["*Der Mathema-*

[35] Heyting en Lakatos, 1967, pág. 173.
[36] Cf. Lakatos, 1963-4, especialmente las págs. 229-35.
[37] J. Myhill, 1967, pág. 175 (subrayado mío); cf. también Lakatos (1963-4).

tik liegt zugrunde"] una sucesión ilimitada de signos o símbolos ["*Zeichen*"] o de sucesiones finitas de símbolos..." [38]. Esto no ha de tomarse como una admisión de la prioridad del lenguaje: sin duda "sucesión" es el término crucial y la idea de sucesión se basa en la intuición del tiempo y en la construcción basada en esta intuición. Sin embargo, muestra que Brouwer era consciente de que los signos o símbolos eran indispensables para llevar a cabo la construcción. Mi propia opinión es que el pensamiento discursivo (es decir, las sucesiones de argumentos lingüísticos) tienen una gran influencia sobre nuestra conciencia del tiempo y sobre el desarrollo de nuestra intuición del orden serial. Esto no contradice en absoluto el constructivismo de Brouwer, pero sí su subjetivismo y mentalismo, ya que ahora los objetos matemáticos pueden pasar a formar parte del tercer mundo: aunque en su origen son construidos por nosotros —el tercer mundo aparece como producto nuestro— los contenidos de pensamiento conllevan sus propias consecuencias involuntarias. La sucesión de los números naturales creada por nosotros, crea los números primos —que *decubrimos*—, los cuales, a su vez, crean problemas que nunca pudimos soñar. *Así es como se hace posible el descubrimiento matemático.* Además, los objetos matemáticos más importantes que descubrimos —los ciudadanos más fecundos del tercer mundo— son *problemas* y nuevos tipos de *argumentos críticos*. Así surge un nuevo tipo de existencia matemática: la existencia de problemas; y un nuevo tipo de intuición: la intuición que nos hace ver problemas y comprenderlos antes de encontrar su solución. (Piénsese en el propio problema central del continuo de Brouwer.)

Heyting ha descrito del modo más esclaracedor la manera en que el lenguaje y el pensamiento discursivo entran en interacción con las construcciones intuitivas más inmediatas (interacción que, dicho sea de paso, destruye ese ideal de certeza evidente y absoluta que la construcción intuitiva creía llevar a cabo). Tal vez deba citar el comienzo de un pasaje de Heyting que no sólo me ha servido de estímulo, sino que además me ha animado: "Lo que en matemáticas es intuitivamente claro ha demostrado no ser claro intuitivamente. Incluso es posible construir una escala descendente de grados de evidencia. El grado superior es el de afirmaciones tales como $2 + 2 = 4$. Sin embargo, $1.002 + 2 = 1.004$ pertenece a un grado inferior, ya que no se pone de manifiesto contando efectivamente, sino que es el razonamiento quien muestra que en general $(n+2) + 2 = n+4$... [Enunciados como éstos] ya poseen el carácter de una implicación: 'Si se construye un número natural n, entonces podemos efectuar la construcción expresada por $(n + 2) + 2 = n + 4$'" [39]. En el contexto en que ahora nos movemos, los "grados de evidencia" de Heyting tienen un interés secundario. Lo que es importante, en primer lugar, es su análisis maravillosamente sencillo y claro de la inevitable interacción entre construcción intuitiva y formulación lingüística que nos sumerge necesariamente en el razonamien-

[38] Brouwer, 1924, pág. 244.
[39] Cf. Heyting, 1962, pág. 195.

to discursivo y, por tanto, lógico. Heyting subraya esta cuestión cuando continúa diciendo: "Este nivel se formaliza en el cálculo de variable libre".

Sobre Brouwer y el platonismo matemático podemos decir una última cosa. La autonomía del tercer mundo es innegable, razón por la cual hemos de rechazar, *al menos* por lo que respecta a los problemas, la ecuación de Brouwer "*esse*=*construi*". Esto puede llevarnos a considerar nuevamente el problema de la lógica del intuicionismo: *sin rechazar las normas de demostración del intuicionismo,* puede ser importante para la discusión racional crítica establecer una separación tajante entre una tesis y los elementos de juicio en su favor. Mas esta distinción se ve destruida por la lógica intuicionista a causa de la *confusión entre los elementos de juicio o demostración y la afirmación a demostrar* [40].

(3') *Problemas metodológicos*: la seguridad constituia el motivo fundamental de la matemática intuicionista de Brouwer, es decir, la búsqueda de métodos de demostración más seguros; de hecho, de métodos infalibles. Ahora bien, si lo que buscamos son pruebas más seguras, debemos ser de lo más riguroso en lo que atañe a la admisión de razonamientos demostrativos: hay que emplear los métodos y suposiciones más débiles. Brouwer se limitó a utilizar medios más débiles que los de la lógica clásica [41]. Demostrar un teorema con medios mas débiles es (y siempre ha sido) una tarea extremadamente interesante y una de las grandes fuentes de problemas matemáticos. De ahí el interés de la metodología intuicionista.

Pero sugiero que esto sólo vale para las demostraciones. Para la crítica, para la refutación, no es deseable una lógica pobre. Aunque el organon de demostración haya de ser débil, el organon de la crítica habrá de ser fuerte. En la crítica, no deseamos limitarnos a demostrar imposibilidades: no exigimos infalibilidad a las críticas, contentándonos a menudo con mostrar que una teoría tiene consecuencias anti-intuitivas. En un organon de crítica, la debilidad y la parquedad no son virtudes, pues es un mérito de las teorías mantenerse frente a críticas fuertes. (Por tanto, parece plausible que en el debate crítico —el metadebate— sobre la validez de una construcción intuicionista pueda ser admisible la utilización de toda la lógica clásica.)

7. Subjetivismo en lógica, teoría de la probabilidad y física

En vista de lo que se ha dicho en la sección 5, sobre todo acerca del empirismo, no es de extrañar que el desprecio por el tercer mundo —y consiguientemente la epistemología subjetivista— siga estando tan extendido en el pensamiento contemporáneo. Aunque no haya una conexión con las

---

[40] Cf. la sección 5.4.

[41] Estas consideraciones sólo se aplican a la *lógica* intuicionista que forma parte de la lógica clásica, aunque la matemática intuicionista no forme parte de la matemática clásica. Véase especialmente las consideraciones de Kleene sobre el "principio de Brouwer" en Kleene y Vesley, 1965, pág. 70.

matemáticas brouwerianas, siguen encontrándose muchas veces tendencias subjetivistas en diversas especialidades. Me referiré aquí a algunas de esas tendencias en lógica, teoría de la probabilidad y física.

7.1. *Lógica epistémica*

La lógica epistémica se ocupa de fórmulas tales como "*a* sabe *p*" o "*a* sabe que *p*" y "*a* cree *p*" o "*a* cree que *p*", que se simbolizan normalmente:

"*Kap*" o "*Bap*",

siendo "*K*" y "*B*" respectivamente las relaciones de saber y creer, *a*, el sujeto cognoscente o creyente y *p* la proposición o el estado de la cuestión sabida o creída.

Mi primera tesis de la sección 1 implica que esto nada tiene que ver con el conocimiento científico: el científico, que llamaré "*S*", ni cree ni sabe. ¿Qué hace? Enumeraré una lista muy breve:

"*S* intenta comprender *p*".
"*S* intenta ingeniar alternativas a *p*".
"*S* intenta ingeniar críticas a *p*".
"*S* propone una contrastación experimental de *p*".
"*S* intenta axiomatizar *p*".
"*S* intenta derivar *p* a partir de *q*".
"*S* intenta mostrar que *p* no es derivable de *q*".
"*S* propone un nuevo problema *x* que surge de *p*".
"*S* propone una nueva solución del problema *x* que surge de *p*".
"*S* critica esta última solución del problema *x*".

Esta lista puede prolongarse un tanto. Su carácter es bastante distinto del de "*S* sabe *p*" o "*S* cree *p*" e incluso del de "*S* cree equivocadamente *p*" o "*S* duda de *p*". De hecho, es algo muy importante que podamos dudar sin criticar y criticar sin dudar. (Que puede ocurrirnos esto es algo que ya vio Poincaré, en su libro *La Ciencia y la Hipótesis* [trad. cast de Alfredo B. Besio y José Banfi; Madrid, Espasa Calpe, Col. Austral, 1943, 3.ª ed. 1963]), que se puede comparar a este respecto con el libro de Russell *Conocimiento del Mundo Exterior* [trad. cast. de M. T. Cárdenas, B. Aires, Cía. General Fabri Editora, 1964.]

7.2. *Teoría de la probabilidad*

En ninguna parte ha encontrado la epistemología subjetivista un apoyo más fuerte que en el campo del cálculo de probabilidades. Este cálculo es una generalización del álgebra de Boole (y por tanto, de la lógica proposicional). Continúa siendo muy corriente interpretarlo en sentido subjetivo como *cálculo de la ignorancia o de la incertidumbre del conocimiento subje-*

tivo. Sin embargo, esto significa interpretar el álgebra de Boole, e incluso el cálculo proposicional, como un cálculo de *cierto conocimiento* —de cierto conocimiento en sentido subjetivo. Pocos bayesianos (como gustan llamarse ahora los partidarios de la interpretación subjetiva del cálculo de probabilidades) aceptarán estas consecuencias.

Durante treinta y tres años he combatido esta interpretación subjetiva del cálculo de probabilidades que surge fundamentalmente de la misma filosofía epistémica que atribuye al enunciado "sé que la nieve es blanca" una dignidad epistémica superior a la del enunciado "la nieve es blanca".

No veo por qué razón no hemos de atribuir una dignidad epistémica aún mayor al enunciado "A la luz de todos los elementos de juicio que tengo a mi disposición estimo que es racional creer que la nieve es blanca". Naturalmente, podría hacerse lo mismo con los enunciados probabilitarios.

7.3. *Física*

El enfoque subjetivo ha progresado mucho en la ciencia a partir de 1926 aproximadamente. Surgió por primera vez en la mecánica cuántica, en la que alcanzó tal potencia que sus oponentes eran considerados como incapaces a los que había que tapar la boca con todo derecho. A continuación, tomó posesión en la mecánica estadística. Szilard propuso en 1929 el punto de vista, ahora aceptado de modo casi universal, según el cual hemos de pagar la información subjetiva con un aumento de entropía física; lo cual se interpretó como demostración de que la entropía física es una carencia de conocimiento y, por tanto, un concepto subjetivo, siendo el conocimiento o información equivalente a la entropía negativa física. Este desarrollo se vio netamente conectado con el desarrollo paralelo de la teoría de la información que, habiendo comenzado como una teoría totalmente objetiva de los canales de comunicación, se vio más tarde ligada al concepto subjetivista de la información de Szilard.

Por tanto, la teoría subjetiva del conocimiento ha invadido la ciencia en un amplio frente. El primer punto de ataque fue la teoría subjetiva de la probabilidad. Mas el virus se ha extendido a la mecánica estadística, la teoría de la entropía, la mecánica cuántica y la teoría de la información.

Evidentemente, en esta conferencia no es posible refutar todas estas teorías subjetivistas. Lo único que puedo decir es que durante años las he combatido (la última vez, en 1967). Mas no abrigo ilusiones. Pueden pasar muchos más años antes de que baje la marea (esperada por Bunge, 1967), si es que baja.

Para terminar, sólo expondré dos cosas.

Primero, intentaré mostrar qué aspecto ofrece la epistemología o la lógica de la investigación desde un punto de vista objetivista y de qué manera puede arrojar alguna luz sobre la biología de la investigación.

En segundo lugar, intentaré señalar en la última sección de esta conferencia, qué aspecto ofrece la psicología de la investigación desde el mismo punto de vista objetivista.

8. Lógica y biología de la investigación

Desde el punto de vista objetivista, la epistemología se convierte en la teoría del aumento de conocimiento. Se convierte en la teoría de la resolución de problemas o, en otras palabras, de la construcción, discusión crítica, evaluación y contrastación crítica de teorías conjeturales que compiten entre sí.

Ahora creo que, por lo que respecta a las teorías rivales, tal vez sea mejor hablar de su "evaluación", "apreciación" o de la "preferencia" de una de ellas, más bien que de su "aceptación". Las palabras no importan demasiado. Además, hablar de "aceptación" no causa ningún daño, si se tiene presente que toda aceptación es provisional y posee, como la creencia, una significación pasajera y personal más bien que objetiva e impersonal [42].

La evaluación o apreciación de las teorías rivales es previa, en parte, a la contrastación (*a priori*, si se quiere, aunque no en el sentido kantiano del término que significa "válida *a priori*") y, en parte, posterior a la contrastación (*a posteriori*, de nuevo en un sentido que no entraña validez). También es previo a la contrastación el contenido (empírico) de una teoría que se relaciona estrechamente con su capacidad explicativa (virtual); es decir, su capacidad de resolver problemas pre-existentes —los que han dado lugar al surgimiento de la teoría y respecto a los cuales *compiten las teorías*.

Sólo en relación a un conjunto pre-existente de problemas pueden evaluarse (*a priori*) las teorías y compararse sus valores. También su llamada simplicidad se puede comparar sólo con respecto a los problemas por cuya solución compiten.

El contenido y el poder explicativo (virtual) son las ideas reguladoras más importantes para la apreciación *apriórica* de las teorías. Se relaciona íntimamente con su grado de contrastabilidad.

La idea más importante para su apreación *a posteriori* es la de verdad o, ya que precisamos un concepto comparador más accesible, lo que he denominado "proximidad a la verdad" o "verosimilitud" [43]. Es importante el hecho de que, aunque pueda ser verdadera una teoría sin contenido (como una tautología), la verosimilitud se basa en la idea reguladora de contenido; es decir, en la idea de la cantidad de consecuencias verdaderas importantes e interesantes de una teoría. Por tanto, una tautología, aunque sea verdadera, tiene un *contenido de verdad* y una verosimilitud cero. Naturalmente, tiene una probabilidad *uno*. Hablando en general, el contenido, la contrastabilidad y la verosimilitud [44] se pueden medir mediante la *improbabilidad*.

---

[42] Por ejemplo, no tengo nada que objetar al uso que hace Lakatos de los términos "aceptación$_1$" y "aceptación$_2$" en su "Changes in the Problem of Inductive Logic", § 3 (Lakatos, 1968).

[43] Cf. Popper 1963, especialmente el capítulo 10, sección 3 y el apéndice 6: también Popper 1962a, en especial la pág. 292; véase también las págs. 29-36 *supra*.

[44] Cf. Popper, "A theorem on truth content", en Feyerabend y Maxwell, 1966.

La evaluación *a posteriori* de una teoría depende totalmente del modo en que ha superado las contrastaciones rigurosas e ingeniosas. Pero las contrastaciones rigurosas, a su vez, presuponen un alto grado de contrastabilidad o contenido *a priori*. Por tanto, la evaluación *a posteriori* de una teoría depende, en gran medida, de su valor *a priori*: las teorías que *a priori* carecen de interés —o tiene escaso contenido— no precisan ser contrastadas, ya que su bajo grado de contrastabilidad excluye *a priori* la posibilidad de someterlas a contrastaciones realmente significativas e interesantes.

Por otro lado, las teorías altamente contrastables son interesantes e importantes aunque no consigan salir indemnes de sus contrastaciones; de su fracaso podemos aprender muchísimo. El fracaso puede ser fructífero, ya que puede sugerir efectivamente cómo construir una teoría mejor.

Sin embargo, en última instancia, tal vez se pueda interpretar el hincapié que hacemos en la importancia fundamental de la evaluación *a priori* como algo debido a nuestro interés por valores *a posteriori* elevados, por obtener teorías que posean un contenido de verdad y verosimilitud elevados, aunque naturalmente sigan siendo siempre conjeturales, hipotéticas o provisionales. Nuestra meta es conseguir teorías que sean no sólo interesantes intelectualmente y altamente contrastables, sino también que, de hecho, hayan salido mejor paradas que sus rivales de las contrastaciones rigurosas; que, por tanto, resuelvan mejor sus problemas y que den lugar a nuevos problemas inesperados y fecundos, si su carácter conjetural se pusiese de manifiesto mediante su refutación.

Podemos decir, por tanto, que la ciencia comienza con problemas y prosigue mediante teorías rivales evaluadas *críticamente*. La evaluación de su verosimilitud posee un significado especial que exige contrastaciones críticas, presuponiendo por tanto, grados de contrastabilidad elevados que, al depender del contenido de la teoría, pueden evaluarse *a priori*.

En la mayoría de los casos, y en los más interesantes, la teoría acabará derrumbándose, dando lugar así a nuevos problemas. El avance conseguido se puede percibir constatando el abismo intelectual existente entre el problema original y el nuevo que ha resultado del derrumbamiento de la teoría.

Este ciclo podemos describirlo mediante el diagrama tan repetido:

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$$

es decir: problema P_1 —teoría provisional— eliminación evaluativa de errores —Problema P_2.

La evaluación es siempre *crítica* y tiene por fin el descubrimiento y la *eliminación de errores*. El aumento del conocimiento —o el proceso de aprendizaje— no es un proceso acumulativo, sino de eliminación de errores. Se trata de una selección darwinista más bien que de una instrucción lamarckiana.

He aquí una breve descripción de la epistemología desde un punto de vista objetivo: el método, o lógica, consistente en tender al aumento del conocimiento objetivo. Mas, aunque se refiera al desarrollo del tercer mundo,

puede interpretarse como una descripción de la evolución biológica. Los animales, e incluso las plantas, son resolutores de problemas. Ademas, resuelven sus problemas mediante el método de intentos competitivos de solución y eliminación de errores.

Las tentativas de solución que incorporan los animales y plantas a su anatomía y comportamiento son biológicamente análogas a las teorías; y viceversa: las teorías se corresponden (igual que muchos productos exosomáticos, como los panales y, especialmente, las herramientas exosomáticas, como las telas de araña) con los órganos endosomáticos y sus modos de funcionamiento. Como las teorías, los órganos y sus funciones son intentos de adaptación al mundo en que vivimos. Además, los nuevos órganos, sus funciones y también los nuevos tipos de comportamiento ejercen su influencia, como las teorías y herramientas, sobre el primer mundo que pueden cambiar. (Un nuevo intento de solución —una teoría, un órgano, un nuevo tipo de comportamiento— puede descubrir un nuevo nicho ecológico virtual actualizándolo.) Un nuevo comportamiento o unos nuevos órganos pueden también llevar al planteamiento de problemas nuevos y de este modo pueden influir en el ulterior curso de la evolución, provocando incluso la emergencia de valores biológicos nuevos.

Todo esto se aplica también a los órganos de los sentidos que llevan incorporadas expectativas semejantes a las teorías. Los órganos de los sentidos, como el ojo, están dispuestos para reaccionar a ciertos acontecimientos seleccionados del medio —a aquellos que "esperan", y *sólo* a ellos—. Como las teorías (y los prejuicios), serán por lo general ciegos a otros acontecimientos que no entienden o no pueden interpretar (porque no corresponden a ningún problema específico que el organismo trate de resolver) [48].

Hay que considerar pre-darwinista a la epistemología clásica que considera "dadas" nuestras percepciones de los sentidos como si fuesen "datos" a partir de los cuales se construyen nuestras teorías mediante un proceso de inducción. No es capaz de darse cuenta del hecho de que los supuestos datos no son más que reacciones adaptativas y, por tanto, interpretaciones que suponen teorías y prejuicios y que, como las teorías, están impregnadas de expectativas de carácter conjeturado; que no puede haber percepción pura, datos puros, del mismo modo que no puede haber un lenguaje observacional puro, desde el momento en que todo lenguaje está impregnado de teorías y mitos. Del mismo modo que nuestros ojos son ciegos a lo imprevisto o inesperado, así nuestros lenguajes son incapaces de describirlo (si bien nuestros lenguajes pueden desarrollarse del mismo modo que pueden hacerlo nuestros órganos de los sentidos, tanto endosomática como exosomáticamente).

La consideración del hecho de que las teorías o expectivas se construyen en nuestros mismos órganos de los sentidos muestra que la epistemología de la inducción se derrumba incluso antes de dar el primer paso. No puede partir de datos de los sentidos o percepciones para construir teorías acerca

[48] Cf. mis consideraciones en Lakatos y Musgrave, 1968, pág. 163.

de ellos, ya que no existen datos de los sentidos o percepciones que no se hayan construido mediante teorías (o expectativas —es decir, los antecesores biológicos de las teorías formuladas lingüísticamente—). Así, los "datos" no son ni base ni garantía de las teorías: no son más seguros que esas teorías o "prejuicios"; en todo caso, menos aún (suponiendo, en aras del argumento, que los datos de los sentidos no sean un invento de los filósofos y existan). Los órganos de los sentidos llevan incorporados el equivalente de teorías primitivas aceptadas acríticamente y contrastadas con menos amplitud que las científicas. Además, no hay lenguaje libre de teoría con el cual describir los datos, ya que los mitos (es decir, las teorías primitivas) surgen junto con el lenguaje. No hay cosas vivas, ni animales ni plantas, sin problemas e intentos de solución equivalentes a las teorías, si bien puede haber vida perfectamente, o así lo parece, sin datos de los sentidos (al menos en las plantas).

Así, la vida, como la investigación científica, va de viejos problemas al descubrimiento de nuevos problemas jamás soñados. Además, este proceso —el de invención y selección— contiene en sí mismo una teoría racional de la emergencia. Los saltos emergentes que conducen a un nuevo nivel son, en primera instancia, los nuevos problemas (P_2) creados por eliminación de errores (*EE*) de una tentativa de solución teórica (*TT*) a un viejo problema (P_1).

9. Descubrimiento, humanismo y autotrascendencia

Para un humanista puede ser interesante nuestro enfoque, ya que sugiere un nuevo modo de considerar la relación entre nosotros —los sujetos— y el objeto de nuestro esfuerzo: el aumento del conocimiento objetivo, el aumento del tercer mundo.

El viejo enfoque subjetivo que interpretaba el conocimiento como una relación entre la mente subjetiva y el objeto conocido —relación que Russell llamaba "creencia" o "juicio"— consideraba lo que yo tengo por conocimiento objetivo como meras *expresiones o manifestaciones* lingüísticas de nuestros estados mentales (o como la conducta correspondiente). Podemos describir este enfoque como *expresionismo epistemológico*, ya que guarda un estrecho parelelismo con la teoría expresionista del arte. La obra de un autor se considera la expresión de sus estados internos: se pone todo el énfasis en la relación causal y en el hecho admitido, aunque pasado por alto, de que el mundo del conocimiento objetivo, como el de la pintura o la música, es una creación del hombre.

Este punto de vista ha de ser sustituido por otro muy distinto. Hay que admitir que el tercer mundo, el mundo del conocimiento objetivo (o, más en general, del espíritu objetivo) es un producto del hombre. Mas hay que subrayar que este mundo posee una existencia en gran medida autónoma que genera sus propios problemas, especialmente los relativos a los métodos de desarrollo, y cuyo impacto sobre cada uno de nosotros, incluso sobre el

pensador creativo más original, escede con mucho al impacto que cualquiera de nosotros pueda tener sobre él.

Sin embargo, sería un error dejar ahí las cosas. El punto para mí más importante no es el simple carácter autónomo y anónimo del tercer mundo o el punto obviamente importantísimo de que siempre se lo debemos casi todo a nuestros predecesores y a la tradición creada por ellos: que, por tanto, debemos al tercer mundo especialmente nuestra racionalidad —es decir, nuestra mente subjetiva, la práctica de los modos de pensar críticos y autocríticos y las disposiciones correspondientes—. Sugiero que más importante que todo esto es la relación entre nosotros mismos y nuestras obras y lo que podemos obtener de esta relación.

El expresionista cree que lo único que puede hacer es dejar que su talento, sus dotes se manifiesten en su obra. El resultado será bueno o malo según el estado mental o fisiológico del trabajador.

Frente a esto sugiero que todo depende del intercambio entre nosotros y nuestra obra, del producto con el que contribuimos al tercer mundo y de la retroalimentación constante que puede ampliarse mediante una autocrítica consciente. Lo increíble de la vida, la evolución y el desarrollo mental es precisamente este método de toma y daca, esta interacción entre nuestras acciones y sus resultados mediante la que nos superamos continuamente y superamos nuestros talentos y dotes.

La autotrascendencia es el hecho más sorprendente e importante de toda la vida y evolución, en especial de la evolución humana.

Evidentemente, la autotrascendencia es menos patente en los estadios pre-humanos, por lo cual puede ser tomada erróneamente por algo así como una autoexpresión. Pero a nivel humano, la autotrascendencia sólo se puede pasar por alto mediante un esfuerzo real. Con nuestras teorías ocurre lo mismo que con nuestros niños: tienden a hacerse cada vez más independientes de sus padres. Además, con nuestras teorías puede ocurrir lo mismo que con nuestros niños: podemos obtener de ellos una cantidad de conocimiento superior a la que les impartimos en un primer momento.

El proceso de aprendizaje, de aumento de conocimiento subjetivo es siempre fundamentalmente el mismo. Consiste en *crítica imaginativa*. Es así como trascendemos nuestro medio local y temporal intentando ingeniar circunstancias *más allá* de nuestra experiencia: criticando la universidad o necesidad estructural de lo que puede parecernos (o lo que los filósofos pueden describir como) lo "dado" o lo "habitual", intentando encontrar, construir, inventar situaciones nuevas —es decir, situaciones *contrastadoras, críticas*— e intentando localizar, detectar y desafiar nuestros prejuicios y suposiciones habituales.

Es así como nos libramos de nuestros propios grilletes, abandonando la ciénaga de nuestra ignorancia; es así como arrojamos una cuerda al aire para subir por ella —si es que logra engancharse en una ramita, aunque sea precariamente.

Lo que hace que nuestros esfuerzos difieran de los de un animal o de los

de una ameba es únicamente que nuestra cuerda puede conseguir un asidero en el tercer mundo de la discusión crítica: el mundo del lenguaje, del conocimiento objetivo. Esto nos permite descartar alguna de nuestras teorías competidoras. Por tanto, si tenemos suerte, podemos lograr sobrevivir a algunas de nuestras teorías equivocadas (y la mayoría de ellas lo son), mientras que la ameba perecerá con sus teorías, creencias y hábitos.

A la luz de todo esto, la vida es resolución de problemas y descubrimiento —descubrimiento de hechos nuevos, de nuevas posibilidades mediante el ensayo de posibilidades tramadas en nuestra imaginación—. A nivel humano, este ensayo se realiza casi completamente en el tercer mundo mediante intentos de representar cada vez con más éxito nuestro primer mundo, y tal vez el segundo, en las teorías de este tercer mundo; y mediante intentos de aproximación mayor a la verdad —una verdad más plena, completa, interesante, lógicamente más fuerte y más relevante— relevante para nuestros problemas.

Lo que se puede llamar segundo mundo —el mundo de la mente— se convierte cada vez más en el nexo entre el primer y el tercer mundos: todas nuestras acciones en el primer mundo están influidas por nuestra captación segundo-mundana del tercer mundo. Por eso es imposible comprender la mente humana y el ego humano sin comprender el tercer mundo (la "mente" o "espíritu objetivo"); y por eso es imposible interpretar sea el tercer mundo como mera expresión del segundo, sea el segundo como mero reflejo del tercero.

Hay tres sentidos del verbo "aprender" que no han sido distinguidos suficientemente por parte de los teóricos del aprendizaje: "descubrir", "imitar", "hacer habitual". Todos ellos pueden considerarse formas de descubrimiento y operan con métodos de ensayo y error que contienen un elemento de azar (que no es demasiado importante y que normalmente se pasa por alto). "Hacer habitual" contiene un mínimo de descubrimiento, si bien nos deja en disposición de realizar ulteriores descubrimientos, con lo que su carácter aparentemente repititivo es falso.

El método de todos estos diversos modos de aprender, adquirir o producir conocimiento es darwinista más bien que lamarckiano: es un problema de selección y no de instrucción repetitiva. (Sin embargo, hemos de reconocer que el lamarckismo es una especie de aproximación al darwinismo y que, por tanto, los productos de la selección suelen mostrarse como si fuesen productos de la adaptación lamarckiana, de la instrucción por repetición: podemos decir que el darwinismo simula el lamarckismo.) Pero la selección es un arma de dos filos: no es sólo el medio quien nos selecciona y cambia; también nosotros seleccionamos y cambiamos el medio, sobre todo descubriendo nuevos nichos ecológicos. A nivel humano, esto lo realizamos en cooperación con un mundo objetivo totalmente nuevo —el tercer mundo, el mundo del conocimiento objetivo provisional que supone nuevas metas y valores tentativos—. No amoldamos o "instruímos" este mundo expresando en él el estado de nuestra mente, ni nos instruye él a nosotros. Tanto noso-

tros como el tercer mundo crecemos ambos mediante lucha mutua y selección. Al parecer, esto ocurre también con las enzimas y los genes: podemos conjeturar que el código genético opera por selección o rechazo más bien que por instrucción o mandato. También parece ocurrir a todos los niveles, hasta con el lenguaje articulado y crítico de nuestras teorías.

Para explicarlo mejor, hemos de considerar los sistemas orgánicos como resultados o productos objetivos del comportamiento tentativo "libre" —es decir, no determinado— dentro de un cierto dominio o rango circunscrito o limitado por su situación interna (especialmente, su composición genética) y su situación externa (el medio). Lo que lleva a la fijación relativa del modo eficaz de reacionar es el fracaso más bien que el éxito. Podemos suponer que el código genético guía la síntesis de las proteínas por el mismo método: impidiendo o eliminando determinadas síntesis químicas posibles más bien que estimulando o guiando directamente. De este modo se puede entender la invención del código genético mediante selección. Convertiría en prohibiciones, resultado de la supresión de errores, las aparentes instrucciones. Como las teorías, el código genético no sólo sería resultado de la selección, sino que además operaría mediante selección, prohibición o impedimento. Naturalmente, se trata de una conjetura, pero es una conjetura atractiva.

BIBLIOGRAFIA SELECTA

ARISTÓTELES, *Metafísica.*

— *De Anima.*

BERKELEY, *Theree Dialogues Between Hylas and Philonous, en Works,* eds. Luce y Jessop, vol. II, 1949. [Versión castellana de Vicente Viqueira, Madrid, Calpe, 1923 (reeditado en la colección Austral de Espasa-Calpe). Hay también una traducción de A. P. Masegosa, Madrid, Aguilar, 1967.]

BOLZANO, B.: *Wissenschaftslehre,* 1837.

BROUWER, L. E. J.: Inaugural lecture, 14 october 1912; trad. de A. Dresden, *Bull. Am. Math. Soc. 20,* 1914, 81-96.

— *Math. Ann. 93,* 1924.

— Mathematik, Wissenschaft und Sprache [Vortrag gehalten in Wien am 10. III, 1928], en *Monatsh. Math. Phys. 36,* 1929, 353-64.

— Consciousness, Philosophy, and Mathematics, en *Proc. 10th Intern. Congress of Philosophy,* 1949, vol. I, fascículo II.

— On Order in the Continuum, and the Relation of Truth to Non-Contradictority, *Kon. Ned. Acad. Recht. Wet. Proc.* Sect. Sci. *54,* 1951.

BÜHLER, K.: *Sprachtheorie,* 1934. [Versión castellana de Julián Marías, *Teoría del Lenguaje,* Madrid, Revista de Occidente, 1967.]

BUNGE, M.: *Quantum Theory and Reality,* 1967.

DESCARTES, R.: *Discourse de la méthode,* 1967; trad. inglesa de E. S. Haldane y G. R. T. Ross, vol. I, 1931. [Hay diversas ediciones castellanas, como la de José Bergua, Madrid, Ediciones Ibéricas; traducción de Antonio Rodríguez Huescar, Madrid, Aguilar, 1954; así como en España-Calpe, Revista de Occidente, etc.]

DUCASSE, C. J.: *J. Phil. 37,* 1940.

FEYERABEND y MAWELL: eds., *Mind, Matter, and Method, Essays in Philosophy and Science in Honor of Herbert Feigl,* 1966.

FREGE, F.: Ueber Sinn und Bedeutung, *Z. Phil. und phil. Kritik, 100,* 1892, 25-50. [Traducción castellana de Ulises Moulines, "Sobre Sentido y Referencia" en Frege, G. *Estudios Sobre Semántica,* Barcelona, Ariel, 1971, págs. 49-84. También, "Sobre Sentido y Denotación", traducción de Alfonso Gómez-Lobo en *Diálogos,* año VIII, núm. 22, abril de 1972; págs. 147-70.]

— Review of Husserl, 1891, *Z. Phil. und phil. Kritik, 103,* 1894, 313-32.

— Der Gedanke. *Beiträge zur Philosophie d. deutschen Idealismus, 1,* 1918.

GOMBRICH, E. H.: Moment and Movement in Art, *J. Warbug and Court. Inst. 27,* 1964.

GOMPERZ, H.: *Weltanschauungslehre,* vol. II, 1, 1908.

— *Über Sinn und Sinngebilde, Verstehen und Erkennen,* 1929.

HAYEK, F. A.: *The Constitution of Liberty,* 1960.

— *Studies in Philosophy, Politics, and Economics,* 1967.

HEGEL, G. W. F.: *Enzyklopädie der philosophischen Wissenschaften,* tercera edición, 1830. [Traducción de Eduardo Ovejero Maury, Madrid, Victoriano Suárez, Comenzada a publicar en 1917.]

HEINEMANN, F.: *Plotin,* 1921.

HENRY, P.: Plotinus' Place in the History of Thought, en *Plotinus, The Enneads,* trad. de S. MacKenna, segunda edición, 1956.

HEYTING, A.: After thirty years, en *Logic, Methodology and Philosophy of Science,* eds, E. Nagel, P. Suppes y A. Tarski, 1962, págs. 194 y sigs.

— *Instuitionism.* 1966. [Traducción de Víctor Sánchez de Zavala, Madrid, Tecnos (en preparación).]

— Informal rigour and intuitionism, en LAKATOS, 1967.

HUSSERL, E.: *Philosophie der Arithmetik,* 1891.

— *Logische Untersuchungen,* vol. I, segunda edición, 1913. [Traducción castellana de José Gaos y Manuel G. Morente, *Investigaciones lógicas.* Madrid, Revista de Occidente, 1929 (reeditado en la colección "Selecta" en 1967).]

KANT, I.: *Kritik der reinen Vernunft,* primera edición, 1770, segunda edición, 1778. [Traducción castellana de José del Perojo. Buenos Aires, Ed. Sopena argentina, 5.ª ed., 1961; trad., de J. del Perojo, revisada por Ansgar Klein. Buenos Aires, Losada, 1967.]

KLEENE, S. C., y R. VESLEY: *The Foundations of Intuitionistic Mathematics.* Amsterdam, North-Holland Publ. Co., 1965.

LAKATOS, I.: Proofs and Refutations, *Brit. J. Phil. of Sci. 14,* 1963-4.

— Editor, *Problems in the Philosophy of Mathematics.* Amsterdam, North-Holland Publ. Co., 1967.

— Editor, *The Problem of Inductive Logic.* Amsterdam, North-Holland Publ. Co., 1968.

— A. MUSGRAVE, eds., *Problems in the Philosophy of Science.* Amsterdam, North-Holland, Publ. Co., 1968.

MYHILL, J.: Remarks on Continuity and The Thinking Subject, en LAKATOS, 1967.

PLATÓN: *Fedón.*

PLOTINO: *Enneades,* ed. Volkmann, 1883, 1884. [Traducción castellana de José Antonio Míguez, *Eneadas.* Madrid, Aguilar, 1963 y siguientes.]

POPPER, K. R.: *Logik der Forschung,* 1934, *The Logic of Scientific Discovery,* 1959 y ediciones posteriores. [Traducción castellana de Víctor Sánchez de Zavala, *La lógica de la investigación científica.* Madrid, Tecnos, 1962.]

— *The Poverty of Historicism,* segunda edición, 1960. [Traducción castellana de Pedro Schwartz, *La Miseria del Historicismo.* Madrid, Haurus, 1961. Reeditada en Alianza Editorial en 1973.]

— *The Open Society and its Enemies,* cuarta edición, 1962 y ediciones posteriores. [Traducción castellana de Eduardo Loedel, *La Sociedad Abierta y sus Enemigos.* Buenos Aires, Paidos, 1957 (reedición en 1967).]

POPPER, K. R.: Some Comments on Truth and the Growth of Knowledge, en *Logic, Methodology and Philosophy of Science*, eds. E. Nagel, P. Suppes y A Tarski, 1962a.
— *Conjectures and Refutations*, 1963 y ediciones posteriores. [Traducción castellana de Néstor Míguez, *El Desarrollo del Conocimiento Científico. Conjeturas y Refutaciones*. Buenos Aires, Paidos, 1967.]
— *Of Clouds and Clocks*, 1966. (Ahora en este volumen, capítulo 6, más adelante ["Sobre Nubes y Relojes"].)
— Quantum Mechanics Without "The Observer", en *Quantum Theory and Reality*, ed. Mario Bunge, 1967.
— On the Theory of Objective Mind, en *Akten des* XIV. *Internationalen Kongresses für Philosophie in Wien*, vol. I, 1968. (Ahora, cap. 4 de este volumen. ["Sobre la Teoría de la Mente Objetiva"].)
— A Pluralist Approach to the Philosophy of History, en *Roads to Freedom, Essays in Honour of Friedrich A. von Hayek*, 1969, págs. 181 y siguientes.
— Eine objektive Theorie des historischen Verstehens, *Schweizer Monatshefte, 50*, 1970, págs. 207 y siguientes.
RUSSELL, B.: On the Nature of Truth, en *Aristotelian Soc. Proc.*, 7, 1906-7, págs. 28-49.
— *Philosophical Essays*, 1910. [Traducción castellana de Juan Ramón Capella, *Ensayos filosóficos*. Madrid, Alianza Editorial, 1968.]
— Introduction to WITTGENSTEIN's *Tractatus*, 1922. [Véase Wittgenstein.]
— *My Philosophical Development*, 1959. [Traducción castellana de Juan Novella Domingo, *La evolución de mi pensamiento filosófico*. Madrid, Aguilar, segunda edición, 1964.]
WATKINS, J. W. N.: *Hobbes's System of Ideas*, 1965.
WHORF, B. L.: *Language, Thought and Reality*, 1956. [Tradución castellana de J. M. Pomares, *Lenguaje, Pensamiento y Realidad*. Barcelona, Barral, 1971.]
WITTGENSTEIN, L.: *Tractatus Logico-Philosophicus*, 1922. [Traducción castellana de E. Tierno Galván. Madrid, Revista de Occidente, 1956 (reedición en Alianza Editorial, 1973).]

4. SOBRE LA TEORIA DE LA MENTE OBJETIVA*

Nuestra tarea fundamental como filósofos consiste en enriquecer nuestra imagen del mundo contribuyendo a la construcción de teorías imaginativas, a la vez que razonadas y críticas, de ser posible, con interés metodológico. La filosofía occidental consiste fundamentalmente en representaciones del mundo que no son sino variaciones sobre el tema del dualismo del cuerpo y la mente, así como en problemas de método relacionados con ellas. Por tanto, las desviaciones fundamentales de esta tradición occidental dualista no han sido más que intentos de sustituir el dualismo por algún tipo de monismo. No me parece que hayan tenido éxito, pues tras el velo de las declaraciones monistas sigue ocultándose el dualismo del cuerpo y la mente.

1. Pluralismo y la tesis de los tres mundos

No obstante, no sólo han aparecido desviaciones monistas, sino también *pluralistas,* lo cual es obvio si pensamos en el politeismo e, incluso, en sus variaciones monoteistas. Con todo, el filósofo abrigará sus dudas acerca de si las diversas interpretaciones religiosas del mundo ofrecen una alternativa auténtica al dualismo del cuerpo y la mente. Los dioses, pocos o muchos, o bien son mentes dotadas de cuerpos inmortales o bien son puras mentes a diferencia de nosotros.

Sin embargo, algunos filósofos han dado algunos pasos serios hacia un pluralismo filosófico, señalando la existencia de un *tercer mundo.* Estoy pensando en Platón, los estoicos y algunos pensadores modernos como Leibniz, Bolzano y Frege, aunque no en Hegel que poseía una fuerte inclinación monista.

En muchos aspectos, el mundo platónico de las formas o ideas era de carácter religioso, pues se trataba de un mundo de realidades superiores. Sin embargo, no era un mundo habitado por dioses personales ni por con-

---

* Conferencia pronunciada (en versión alemana abreviada) el 3 de septiembre de 1968 en Viena. Reproducida a partir de los *Akten des XIV. Internationalen Kongresses für Philosophie*, vol. I, Viena 1968, págs. 25-53. He añadido ahora algunas cuestiones que se publicaron originalmente (en alemán) en *Schweizer Monatshefte*, 50. Jahr, Heft 3, 1970, págs. 207-15. [El presente artículo —ahora ligeramente modificado— constituyó la aportación del autor al Simposium sobre "Filosofía de la ciencia", celebrado en Burgos del 23 al 25 de septiembre de 1968 en homenaje a K. R. Popper. Cfr. "Sobre la Teoría de la inteligencia objetiva", trad. cast. de Víctor Sánchez de Zavala en *Simposio de Burgos. Ensayos de Filosofía de la Ciencia: en torno a la obra de Sir Karl R. Popper.* Tecnos, Madrid, 1970, págs. 202-237.]

ciencias o contenidos de alguna conciencia. Se trataba de un mundo objetivo y autónomo que existía a parte del mundo físico y del mundo de la mente.

Estoy de acuerdo con los intérpretes de Platón que afirman que sus formas o ideas son diferentes, no sólo de los cuerpos y las mentes, sino también de las "ideas que están en la mente"; es decir, de las experiencias conscientes o inconscientes: las formas o ideas de Platón constituyen un tercer mundo *sui generis*. Hay que admitir que pueden ser *objetos* de pensamiento virtuales o posibles —*intelligibilia*—, aunque Platón los considera tan objetivos como los *visibilia* que son cuerpos físicos: objetos visuales, virtuales o posibles [1].

Por tanto, el platonismo supera la dualidad del cuerpo y la mente al introducir un mundo tripartito o, como prefiero decir, un tercer mundo.

Sin embargo, no voy a discutir aquí sobre Platón, sino sobre el pluralismo. Aunque estemos en un error los que atribuimos a Platón este pluralismo, puedo apelar a una conocida *interpretación* de su teoría de las formas o ideas como ejemplo de filosofía que trasciende efectivamente el esquema dualista.

Voy a partir de la fisolofía pluralista para mi exposición, aunque no soy ni platónico ni hegeliano [2].

Para esta filosofía pluralista, el mundo consta al menos de tres sub-mundos ontológicamente distintos: el primero, es el mundo físico o de los estados físicos; el segundo, es el mundo mental o de los estados mentales; el tercero, es el de los inteligibles o de las *ideas en sentido objetivo*, el mundo de los objetos de pensamiento posibles: el mundo de las teorías en sí mismas y sus relaciones lógicas, de los argumentos y de las situaciones problemáticas tomados en sí mismos.

Una de las dificultades fundamentales de esta filosofía pluralista se refiere a las relaciones entre estos tres "mundos". Están relacionados entre sí de tal modo que tanto los dos primeros como los dos últimos pueden entrar en interacción [3]. Por tanto, el segundo mundo de las experiencias subjetivas o personales mantiene relaciones con los otros dos. El primero y el tercero sólo se relacionan por intermedio del segundo, el mundo de las experiencias personales o subjetivas.

---

[1] Para la distinción platónica entre lo visible (*horaton*) y lo inteligible (*noeton*) véase, por ejemplo, la *República*, 509E. (Cf. el *Teeteto*, 185D y sigs.) La fisiología del ojo ha puesto de manifiesto la estrecha semejanza existente entre los procesos de percepción visual de los *visibilia* y una interpretación elaborada de los *intelligibilia*. (Puede decirse que Kant anticipó gran parte de esto.)

[2] Hegel, siguiendo a Aristóteles, rechazó el tercer mundo platónico. Al confundir los objetos con los procesos de pensamiento, atribuyó desastrosamente la conciencia al espíritu objetivo, deificándolo. (Véase, especialmente, el final de la *Enciclopedia* de Hegel con la oportuna cita de la *Metafísica* de Aristóteles, 1072b 13-30.)

[3] Empleo aquí la palabra "interacción" en sentido amplio, de modo que no excluya un paralelismo psicofísico: no es mi intención discutir aquí este problema. (En otros lugares he defendido el interaccionismo: véanse por ejemplo los capítulos 12 y 13 de mi libro *Conjectures and Refutations*, 1963, 1965, 1969 [traducción castellana de Néstor Míguez, *El Desarrollo del Conocimiento Científico. Conjeturas y Refutaciones*, Buenos Aires, Paidos, 1967; págs. 339-50].)

2. Las relaciones causales entre los tres mundos

Considero de la mayor trascendencia descubrir y explicar de este modo las relaciones entre los tres mundos; es decir, considerando al segundo como mediador entre el primero y el tercero. Aunque rara vez se explicita, me parece evidente que este punto de vista está contenido en la teoría de los tres mundos, según la cual la mente humana puede ver un cuerpo físico en el sentido literal de la palabra que entraña la participación de los ojos en el proceso. La mente también puede "ver" o "captar" un objeto geométrico o aritmético, un número o una figura geométrica. Ahora bien, en este sentido, "ver" o "captar" se usan de modo metafórico, pero de todos modos denotan una relación real entre la mente y el objeto inteligible, el objeto aritmético o geométrico, siendo esta relación en gran medida análoga a la de "ver" en sentido literal. Por tanto, la mente puede enlazar con objetos tanto del primero como del tercer mundo.

Mediante estos lazos, la mente establece un nexo *indirecto* entre los mundos primero y tercero que resulta ser de suma importancia. No se puede negar seriamente que el tercer mundo de las teorías matemáticas y científicas ejerza una gran influencia sobre el primero. Esto se lleva a cabo, por ejemplo, mediante la intervención de los técnicos que introducen modificaciones en el primer mundo, sirviéndose para ello de determinadas consecuencias de las teorías; teorías que, dicho sea de paso, se han desarrollado primitivamente gracias a otras personas inconscientes de las posibilidades tecnológicas inherentes a sus teorías. Por tanto, estas posibilidades estaban ocultas en las teorías mismas, en las ideas objetivas mismas, siendo descubiertas en su seno por personas que intentaron *comprender* esas ideas.

Creo que si lo desarrollamos detenidamente, este argumento apoya la realidad objetiva de los tres mundos. Apoya, además, no sólo la tesis de la existencia de un mundo mental y subjetivo de experiencias personales (cosa que niegan los conductistas), sino también la tesis según la cual una de las funciones fundamentales del segundo mundo consiste en captar los objetos del tercero. Es algo que todos hacemos: es característica esencial del ser humano aprender un lenguaje, lo cual significa básicamente aprender a captar *contenidos de pensamiento objetivos* (como los llamaba Frege) [4].

Creo que algún día habremos de hacer una revolución en psicología al considerar la mente humana como órgano de interacción con los objetos del tercer mundo para comprenderlos, desarrollarlos, utilizarlos y hacerlos repercutir sobre el primer mundo.

---

[4] Cf. Gottlob Frege, "*Über Sinn und Bedeutung*", *Zeitschrift für Philosophie und philosophische Kritik, 100,* (1892), pág. 32 [traducción castellana de Ulises Moulines, "Sobre Sentido y Referencia" en G. Frege, *Estudios sobre Semántica,* Barcelona, Ariel, 1972; pág. 58, nota 5]: "Entiendo por pensamiento, no el acto subjetivo de pensar, sino su contenido objetivo..."

3. La objetividad del tercer mundo

El tercer mundo o, mejor dicho, los objetos que pertenecen a él, las formas o ideas objetivas descubiertas por Platón, se han confundido muy a menudo con las ideas subjetivas o procesos que pertenecen al segundo mundo más bien que al tercero.

Este error tiene su historia que comienza con el propio Platón. Aunque está claro que se dio cuenta del carácter de pertenencia al tercer mundo que poseían sus ideas, parece ser que no se percató de que el tercer mundo no sólo contenía nociones o conceptos universales, como el número 7 o el 77, sino también verdades matemáticas o proposiciones [5], como por ejemplo, "7 por 11 es igual a 77", e incluso proposiciones falsas como "7 por 11 es igual a 66", así como todo tipo de proposiciones o teorías no matemáticas.

Parece que fueron los estoicos los primeros en darse cuenta de la cuestión al desarrollar una filosofía del lenguaje de una sutileza extraordinaria. Se dieron cuenta de que el lenguaje humano pertenece a los tres mundos [6]. Pertenece al primero por cuanto que consta de acciones físicas o símbolos físicos; al segundo, por cuanto que expresa un estado subjetivo o psicológico o por cuanto que captar o comprender un lenguaje entraña un cambio en nuestros estados subjetivos [7]; y, al tercero, por cuanto que el lenguaje contiene información, dice, enuncia o describe algo o comunica un significado o mensaje significativo que implica, concuerda o contradice a otro. *Las teorías, proposiciones o enunciados son las entidades lingüísticas más importantes del tercer mundo.*

Cuando decimos "he visto algo escrito en un papiro" o "he visto algo gravado en bronce", hablamos de entidades lingüísticas pertenecientes al primer mundo: no entramos en el problema de si podemos leer o no el mensaje. Cuando decimos "me impresionó mucho la seriedad y convinción con que pronunció la conferencia" o "eso, más que un enunciado, ha sido un exabrupto airado" hablamos de entidades lingüísticas pertenecientes al

---

[5] En el *Teeteto*, 139E y sig. Platón parece sugerir que la verdad y las proposiciones no son (normalmente) ideas del tercer mundo, sino actos mentales (como lo son los actos de captar las nociones de semejanza, etc., descritos un poco antes (186A). Dice allí: "el pensamiento es el diálogo del alma consigo misma acerca de un objeto cualquiera". Cf. el *Sofista*, 263E-264B, donde se hace hincapié en el habla silenciosa (verdadera o falsa), la afirmación, negación y opinión. Pero, en el *Fedro*, 247D-249B, la verdad es uno de los huéspedes del tercer mundo captados por el alma.

[6] Los estoicos eran materialistas: consideraban el alma como una parte del cuerpo, identificándola con "el aliento vital" (Diógenes Laercio, VII, 156 y sig. [Traducción castellana de José Ortiz y Sanz, Espasa-Calpe (Austral), Buenos Aires, 1950]). Consideraban la facultad de razonar como "el elemento rector" del cuerpo (Sexto, *Adv. Mat.*, VIII, 39 y sigs.). No obstante, esta teoría se puede interpretar como un tipo especial de dualismo cuerpo-mente, ya que presenta una solución especial del problema. Si a estos dos mundos (o partes del primero) añadimos el *contenido de "lo dicho"* (*lecton*) llegamos a la versión estoica del tercer mundo.

[7] Al parecer, la idea de *estado* mental (así como las de bondad o veracidad) pertenecen a los estoicos; naturalmente, se entiende en el sentido de un estado del aliento y, por tanto, del cuerpo. Cf. Sexto, *loc. cit.*

segundo mundo. Cuando decimos "sin embargo, Jaime ha dicho hoy exactamente lo contrario de lo que dijo ayer Juan" o "de lo que dice Jaime se desprende claramente que Juan está equivocado" o cuando hablamos del platonismo o de la teoría cuántica, hablamos de algo objetivo, de un *contenido lógico objetivo;* es decir, hablamos de la significación, en el tercer mundo, de la información o mensaje transmitido por lo que se ha dicho o escrito.

Fueron los estoicos quienes establecieron por primera vez la importante distinción entre el *contenido* lógico objetivo (del tercer mundo) de lo que decimos y los *objetos* de los que hablamos. Estos, a su vez, pueden pertenecer a uno cualquiera de los tres mundos: podemos hablar, en primer lugar, acerca del mundo físico (cosas o estados físicos); en segundo lugar, podemos hablar sobre nuestros estados mentales subjetivos (incluyendo nuestra comprensión de una teoría); en tercer lugar, podemos hablar acerca del contenido de algunas teorías (como puede ser una serie de proposiciones aritméticas) o, pongamos por caso, acerca de su verdad o falsedad.

Considero muy recomendable intentar evitar términos tales como "expresión" y "comunicación" para referirnos a lo que se dice en el sentido del tercer mundo, ya que "expresión" y "comunicación" son términos esencialmente psicológicos, siendo peligrosas sus connotaciones subjetivistas o personales, sobre todo, en un terreno en el que tan fuerte es la tentación de interpretar los contenidos del pensamiento del tercer mundo en términos de procesos de pensamientos del segundo.

Es interesante que, partiendo de ideas platónicas, los estoicos no se limitasen a aplicar la teoría del tercer mundo a las teorías y proposiciones, sino que además de estas entidades lingüísticas del tercer mundo, como son los enunciados declarativos o aserciones, también incluían cosas tales como problemas, argumentos, investigaciones argumentadoras e incluso órdenes, avisos, súplicas, tratados y, naturalmente, la poesía y la narración. También introducían una distinción entre el estado personal de veracidad y la verdad de una teoría o proposición; es decir, una teoría o proposición a la que se aplica el predicado del tercer mundo "objetivamente verdadero".

4. El tercer mundo como producto del hombre

Podemos distinguir dos grupos fundamentales de filósofos. El primero, está formado por aquellos que, como Platón, aceptan un tercer mundo autónomo, considerándolo sobrehumano, divino y eterno. El segundo, lo constituyen aquellos que, como Locke, Mill, Dilthey o Collingwood señalan el carácter de *producto humano* que posee el *lenguaje* junto con todo lo que "expresa" o "comunica" y que, por esta razón, consideran todo lo lingüístico como una parte de los mundos primero y segundo, rechazando cualquier sugerencia acerca de la existencia de un tercero. Es interesante que la mayoría de los que se dedican a las humanidades formen parte del grupo que rechaza el tercer mundo.

El primer grupo, los platónicos, se siente seguro de su postura por el hecho de que podamos hablar de verdades eternas: una proposición es atemporalmente verdadera o falsa. Esta impresión es de importancia decisiva: las verdades eternas tienen que haber sido verdaderas antes de que existiese el hombre, por lo cual no pueden ser producidas por nosotros.

Los miembros del segundo grupo están de acuerdo en que las verdades eternas no pueden ser producidas por nosotros, de donde infieren que *no pueden ser "reales"*: lo que es "real" no es más que el *uso* que *nosotros* hacemos del predicado "verdadero" junto con el hecho de que, al menos en ciertos contextos, lo utilicemos como predicado independiente del tiempo. Pueden argüir que este tipo de utilización no es tan sorprendente como pudiera parecer a primera vista: aunque Pablo, el padre de Pedro, pueda pesar más que éste en un momento dado y menos un año después, nada de eso puede ocurrir con dos trozos de metal en tanto en cuanto uno de ellos continúe siendo una pesa exacta de una libra y el otro, una pesa exacta de dos libras. Aquí, el predicado "exacto" desempeña la misma función que el predicado "verdadero" por lo que respecta a los enunciados; de hecho, podemos sustituir "exacto" por "verdadero". Con todo, pueden decir estos filósofos, nadie negará que las pesas pueden ser hechas por el hombre.

Creo que existe la posibilidad de mantener una postura distinta de la de ambos grupos de filósofos: sugiero la *posibilidad de aceptar la realidad o (como también puede decirse) la autonomía del tercer mundo y, a la vez, admitir que éste se constituye como producto de la actividad humana.* Incluso se puede admitir que el tercer mundo es un producto humano a la vez que sobrehumano en un sentido muy claro [*]. Trasciende a su productor.

Si tenemos en cuenta la pretendida efectividad que puede tener el tercer mundo sobre el primero, sirviéndose del segundo, veremos con toda claridad que no es una ficción, sino que existe "en realidad". Basta pensar en el impacto de la trasmisión de la energía eléctrica o de la teoría atómica sobre nuestro medio orgánico e inorgánico o en el impacto de las teorías económicas sobre las decisiones acerca de si construir un avión o un barco.

Según la postura que he adoptado aquí, el tercer mundo (una de cuyas partes está constituida por el lenguaje humano) es un producto del hombre, del mismo modo qua la miel es producida por las abejas o las telas de araña, por las arañas. Aunque puedan constituir soluciones a problemas biológicos o de otro tipo, grandes regiones del tercer mundo, entre las que se encuentra el lenguaje humano, son, como el lenguaje (y como la miel), *productos de*

---

[*] Aunque sea hecho por el hombre, el tercer mundo (tal como entiendo yo este término) es sobrehumano por cuanto que sus contenidos son objetos de pensamiento virtuales más bien que actuales y en el sentido de que tan sólo pueden convertirse en objetos actuales de pensamiento un número finito de los infinitos objetos virtuales. No obstante, hemos de guardarnos mucho de interpretar estos objetos como pensamientos de una conciencia sobrehumana, como fue el caso, por ejemplo, de Aristóteles, Plotino y Hegel. (Véase la nota primera.) Para el carácter sobrehumano de la verdad, veánse las págs. 29 y sigs. de mi libro *Conjectures and Refutations*, 1963 [trand. cit., págs. 39 y sigs.].

actividades humanas que no están específicamente diseñadas para ello *.

Consideremos la teoría de números. Creo (en contra de Kronecker) que incluso los números naturales son una obra humana, el producto del lenguaje y del pensamiento del hombre, a pesar del hecho de que el número de ellos es infinito; hay más de los que el hombre pueda pronunciar nunca y más de los que puedan utilizarse con las computadoras. Hay, además un número infinito de ecuaciones verdaderas y falsas entre dichos números; muchas más de las que podamos declarar verdaderas o falsas.

Además, lo cual es aún más interasante, surgen problemas nuevos e inesperados como subproducto involuntario de la sucesión de los números naturales, como por ejemplo, los problemas no resueltos de la teoría de los números primos (como la conjetura de Goldbach). Estos problemas son claramente *autónomos*. No son producidos por nosotros en nigún sentido, sino que los *descubrimos* y, en este sentido, existen sin descubrir antes de que demos con ellos. Es más, puede que algunos sean insolubles.

En nuestros intentos por resolver éstos u otros problemas podemos inventar nuevas teorías, las cuales, una vez más, son producidas por nosotros:

---

* Véase la teoría de Karl Bühler acerca de las funciones superiores e inferiores del lenguaje humano y mi desarrollo de la teoría en *Conjectures and Refutations*, 1963, págs. 134 y sigs. y 295 y sigs. [págs. 158 y sigs. y 341 y sigs.], así como en mi artículo *Sobre Nubes y Relojes*, 1966; véanse más adelante las págs. 218-220. Véase también F. A. Hayek, *Studies in Philosophy, Politics and Economics*, 1967, especialmente los capítulos 3, 4 y 6. Resumiendo, Bühler señala que los lenguajes animales y humanos se parecen en cuanto que constan siempre de *expresiones* (síntomas de un estado del organismo) y *comunicaciones* (señales). Sin embargo, el lenguaje humano se distingue por poseer *además* una función superior: puede ser *descriptivo*. Por mi parte, he apuntado otras funciones superiores, en especial una que considero de importancia decisiva: la función *argumentadora o crítica*.

Es importante que esta teoría señale el hecho de que estas funciones inferiores estén siempre presentes. (Por tanto, no le afecta la crítica que hace R. G. Collingwood —en su libro *Principles of Art*, 1938, págs. 262 y sigs.— a la teoría del lenguaje de I. A. Richards, expuestas en *The Principles of Literary Criticism*, 2.ª edición, 1926.)

Por lo que respecta al significado de las consecuencias involuntarias de las acciones humanas intencionales, véase Hayek, *op. cit.*, pág. 100, en especial la nota 12. Por lo que atañe al origen del lenguaje, creo que fue Hayek el primero que llamó mi atención sobre un pasaje del *Discurso del Método*, segunda edición (Haldane y Ross, vol. I, pág. 89 [Edicciones Ibéricas, Madrid, 2.ª parte, pág. 143]) en el que Descartes describe el desarrollo y perfeccionamiento del "camino real" como consecuencia involuntaria de su uso, teoría que puede aplicarse al desarrollo del lenguaje. Me he ocupado con cierta extensión del problema de las consecuencias involuntarias de las acciones intencionadas en mi libro *The Poverty of Historicism*, 1944, 1957, página 65. [Traducción castellana de Pedro Schwartz, *La Miseria del Historicismo*. Madrid, Taurus, 1961, pág. 88, nota 14; reeditada en Alianza Editorial, Madrid, 1973.] (publicado después del libro de Hayek *The Counter Revolution of Science*, 1942, 1952, aunque fue escrito antes de 1942). En una nota, aludo allí a Hume y a "una explicación darwinista del carácter instrumental de las instituciones no proyectadas". También me ocupo de ello en mi libro *The Open Society and its Enemies*, 1945, en especial, vol. II. capítulo 14, págs. 93-8 y nota 11 de las páginas 323 y sigs. [Traducción castellana de Eduardo Loedel, *La Sociedad Abierta y sus Enemigos*, Buenos Aires, Paidos, 1957, págs. 295-8 y 632.] (en su libro *Studies in Philosophy*, pág. 100, nota 12, Hayek ha hecho una crítica de lo que allí digo que tengo que agradecer). Véase también mi conferencia "Epistemología sin Sujeto Cognoscente" (pronunciada en Amsterdam en 1967) reproducida ahora en este volumen en el capítulo 3.

son un resultado de nuestro pensamiento crítico y creador, asistido en gran medida por otras teorías existentes del tercer mundo. Sin embargo, una vez que hemos producido dichas teorías, nos crean problemas nuevos, inesperados y no buscados, problemas autónomos, problemas a descubrir.

Esto explica que el tercer mundo sea *autónomo* por lo que respecta a lo que podríamos denominar su condición ontológica, aunque genéticamente sea producido por nosotros; y también, que podamos actuar sobre él, aportando cosas y facilitando su desarrollo, aunque nadie pueda dominar ni siquiera una pequeña parte de dicho mundo. Todos contribuimos a su desarrollo, si bien todas las contribuciones individuales son insignificantemente pequeñas. Todos tratamos de aprehenderlo, ya que no podemos vivir sin entrar en contacto con él, porque todos hacemos uso del lenguaje, cosa sin la cual difícilmente podríamos ser humanos [10]. Sin embargo, el tercer mundo ha ido creciendo hasta más allá del alcance, no sólo de cualquier hombre, sino también de todos los hombres juntos (como muestra la existencia de problemas insolubles). Sus efectos sobre nosotros se han hecho más importantes para nuestro desarrollo y el suyo propio que los de nuestra acción creadora sobre él, ya que casi la totalidad de su crecimiento se debe a un efecto de retroalimentación: al reto que representa el descubrimiento de problemas autónomos, muchos de los cuales nunca dominaremos. [11] Además, siempre tendremos delante la tarea de descubrir problemas nuevos, ya que hay una infinidad de problemas que siempre quedarán sin descubrir. A pesar —y también a causa— de la autonomía del tercer mundo, siempre habrá lugar para el trabajo creador y original.

5. El problema de la comprensión

He dado aquí algunas razones en favor de la existencia autónoma de un tercer mundo objetivo porque deseo hacer una contribución a *la teoría de la comprensión* (*"hermenéutica"*) que tanto han discutido los estudiosos de las humanidades (*"Geisteswissenschaften"*, "ciencias morales o del espíritu"). Partiré aquí del supuesto de que el problema central de las humanidades es *la comprensión de los objetos que pertenecen al tercer mundo*. Tal vez pueda dar la impresión de que de este modo nos alejamos radicalmente del dogma fundamental aceptado por casi todos los humanistas (como

[10] Helen Keller ha descrito de un modo convincente y emotivo la capacidad humanizadora de su dramático descubrimiento del lenguaje. De todas las funciones del lenguaje especialmente humanizadoras, la argumentadora o crítica es la que considero más importante: es la base de lo que denominamos racionalidad humana.

[11] La causa de ello estriba en que podemos mostrar (A. Tarski, A. Mostowski, R. M. Robinson, *Undecidable Theories*, Amsterdam, 1953; véase especialmente la nota 13 de la pág. 60 y sig.) que el sistema (completo) de todas las proposiciones verdaderas de la aritmética de los enteros no es axiomatizable, siendo esencialmente indecidible. De ahí se sigue que siempre habrá un número infinito de problemas sin resolver en aritmética. Es interesante que podamos llevar a cabo tales descubrimientos inesperados sobre el tercer mundo que son en gran medida independientes de nuestro estado mental. (Este resultado ha de retrotraerse hasta Kurt Gödel, el iniciador de estos trabajos.)

su propio nombre indica) y, en especial, por la mayoría de los que se interesan por el problema de la comprensión. Me refiero, evidentemente, al dogma de que los objetos de nuestra comprensión pertenecen fundamentalmente al segundo mundo o, en todo caso, de que han de ser explicados en términos psicológicos [12].

Es evidente que los procesos o actividades abarcados por el concepto-sombrilla "comprensión" son actividades subjetivas personales o psicológicas. Pero han de distiguirse del *resultado* (más o menos fecundo) de estas actividades, de lo que de ellas se deriva: el "estado final" (por el momento) de la comprensión, de la *interpretación*. Aunque esto *pueda* constituir un estado de comprensión subjetivo, también puede ser un objeto del tercer mundo, especialmente una teoría. Este último caso es, en mi opinión, el más importante. La interpretación, en cuanto objeto del tercer mundo, será siempre una teoría, como por ejemplo, una explicación histórica apoyada por una cadena de razonamientos y quizá por elementos de juicio documentales.

Por tanto, toda interpretación es una especie de *teoría* y, como tal, está engarzada con otras teorías y con otros objetos del tercer mundo. De este modo es como puede plantearse el problema, relativo al tercer mundo,

---

[12] A pesar de la moda antipsicologista iniciada con las *Logische Untersuchungen*, 1900-1 de Husserl (2.ª ed., 1913, 1921) [Traducción castellana de Gaos y Morente, *Investigaciones Lógicas*, Madrid, Revista de Occidente, 1929; reimpresión en "Selecta" en 1967, dos vols.], el psicologismo —es decir, el olvido e incluso la negación del tercer mundo— sigue teniendo una gran vigencia, especialmente entre los que se interesan por la teoría de la comprensión ("hermenéutica"). No cabe duda de que el antipsicologismo de Husserl procede de la crítica de Frege a la psicologista *Philosophie der Arithmetik. Psychologische und Logische Untersuchungen* de Husserl (1891). En sus *Logische Untersuchungen* (donde hace alusión a Bozano) Husserl enuncia con maravillosa claridad (vol. I, pág. 178 [trad. cit., pág. 207 del vol. I]): "En todas... las ciencias hemos de insistir en la distinción fundamental que existe entre los tres tipos de interrelación: (a) las interrelaciones entre nuestras *experiencias cognitivas*..." (se trata de lo que llamo aquí *segundo mundo*) (*b*) "La interrelación entre los *objetos que se investigan*..." (en especial, mi *primer mundo*, aunque también pueden ser los otros) "y (c) Las interrelaciones *lógicas*...", que pertenecen a mi tercer mundo). Sin embargo, pudiera ocurrir que precisamente este pasaje importantísimo hubiese de ser condenado a causa de una confusión aún dominante. La razón de ello es que en el pasaje del apartado (a) señalado aquí con puntos suspensivos, Husserl alude a las interrelaciones psicológicas de "juicios, intelecciones, conjeturas, problemas" y especialmente, también a actos de *comprensión intuitiva* "por lo que una teoría descubierta desde hace tiempo se piensa con penetración intelectiva". La alusión a "juicios", "conjeturas" y "problemas" (al mismo nivel que las "intelecciones") puede haber llevado a confusiones, sobre todo, porque Husserl habla en el apartado (c) *sólo de verdades*, excluyendo aparentemente las proposiciones falsas, las conjeturas, los problemas y los interrogantes: habla de "las *verdades* de una disciplina científica, especialmente de una teoría científica, de una demostración o conclusión". (Hay que recordar que tanto Husserl como muchos otros pensadores más recientes consideraban las teorías científicas como hipótesis científicas cuya *verdad había de ser probada*: la tesis del carácter conjetural de las teorías científicas estaba muy desacreditada por su carácter absurdo cuando traté de propugnarla en los años treinta.) El modo en que Husserl habla de la *comprensión* en este pasaje (cf. también el vol. II, págs. 62 y sigs. [Vol. I, págs. 356 y sigs.]) puede ser también responsable de algunas tendencias psicologistas aún vigentes.

de los *méritos* de la interpretación y, en especial, su valor para nuestra *comprensión* histórica.

Pero incluso el acto subjetivo o el estado disposicional de "*comprender*" puede entenderse, a su vez, sólo mediante sus conexiones con los objetos del tercer mundo, ya que sostengo las tres tesis siguientes acerca del acto subjetivo de compresión.

(1) Todo acto subjetivo de comprensión está engarzado en gran medida con el tercer mundo;

(2) Casi todas las consideraciones que pueden hacerse en torno a dicho acto consisten en señalar sus relaciones con los objetos del tercer mundo, y

(3) Dicho acto consta fundamentalmente de operaciones hechas con objetos del tercer mundo: operamos con ellos casi como si fuesen objetos físicos.

Sugiero que esto se puede generalizar y que se aplica también a todo acto subjetivo de "conocimiento": todas las cosas importantes que podemos decir de un acto de conocimiento consisten en señalar los objetos del acto (una teoría o proposición) pertenecientes al tercer mundo, así como sus relaciones con otros objetos del tercer mundo, tales como los argumentos relevantes para el problema, amén de los objetos conocidos.

6. Procesos psicológicos de pensamiento y objetos del tercer mundo

Me temo que, incluso quienes aceptan la necesidad de analizar el *estado final de la comprensión* (*subjetiva*) en términos de los objetos del tercer mundo, rechacen la tesis correspondiente relativa a *la actividad personal o subjetiva de captar o comprender*: generalmente se cree que no podemos arreglárnoslas sin tales procedimientos subjetivos, *como la comprensión simpatética*, la empatía, la reproducción subjetiva de las acciones de otras personas (Collingwood) o el intento de colocarnos en la situación de otra persona haciendo nuestros sus problemas e ideales.

Frente a ello, sostengo la tesis siguiente. Tanto el estado de comprensión subjetivo alcanzado finalmente, como el proceso psicológico que conduce a él han de ser analizados en términos de los objetos del tercer mundo en que están enraizados. De hecho, *sólo* se puede hacer un análisis en esos términos. El proceso o actividad de comprensión consta fundamentalmente de una sucesión de estados de comprensión. (El que uno de ellos sea o no un estado "final" a menudo puede depender, desde un punto de vista subjetivo, tan sólo de algo tan carente de interés como puede ser un sentimiento de agotamiento.) Sólo se puede afirmar algo más acerca de él, si hemos conseguido dar con un argumento importante o un nuevo elemento de juicio —es decir, un objeto del tercer mundo—. Hasta tanto, lo que constituye el

"proceso" no es más que la sucesión de los estados precedentes y lo que constituye la "actividad" no es otra cosa que el trabajo de criticar el estado alcanzado (es decir, el trabajo de formular argumentos críticos del tercer mundo). Para decirlo con otras palabras: *la actividad de comprender consiste esencialmente en operar con objetos del tercer mundo.*

La actividad se puede representar mediante un *esquema general de resolución de problemas por el método de conjeturas imaginativas seguidas de crítica* o, como tantas veces he dicho ya, *por el método de conjeturas y refutaciones*. El esquema (en su forma más simple) es el siguiente [13]:

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$$

Aquí P_1 es el *problema* del que partimos y *TT* (la "teoría tentativa") la primera solución imaginitiva que ingeniamos; por ejemplo, nuestro primer *intento de interpretación*. *EE* (*"eliminación de errores"*) consiste en el examen crítico riguroso de nuestra conjetura, de nuestra interpretación provisional: consiste, por ejemplo, en la utilización crítica de los documentos que sirven como elementos de juicio y, si en este estadio primitivo disponemos de más de una conjetura, consistirá también en la discusión crítica y comparación evaluativa de las conjeturas rivales. P_2 es la situación problemática tal como surge del intento crítico de resolver nuestros problemas y que lleva a un segundo intento (*y así sucesivamente*). Se logrará una comprensión satisfactoria, si la interpretación, la teoría conjetural, encuentra apoyo en el hecho de poder arrojar nueva luz sobre nuevos problemas —sobre más problemas de los que esperábamos—; o encuentra apoyo en el hecho de explicar muchos subproblemas, algunos de los cuales no veíamos al comienzo. Así podemos decir que es posible evaluar el progreso realizado comparando P_1 con algunos problemas posteriores (digamos, P_n).

Este análisis esquemático tiene una gran aplicabilidad y opera plenamente sobre objetos del tercer mundo, como problemas, conjeturas y argumentos críticos. Con todo, también resulta ser un análisis de lo que hacemos en nuestro segundo mundo subjetivo cuando intentamos comprender algo.

Sin duda, un análisis más detallado nos haría ver que siempre tomamos nuestros problemas del *trasfondo* del tercer mundo [14]. Este trasfondo está

---

[13] Este esquema tetrádico, junto con una versión más elaborada, se puede encontrar en la sección XVIII de *Sobre Nubes y Relojes*, 1966 (que aparece reimpreso en este volumen en el capítulo 6). Puede considerarse el resultado de la interpretación crítica del esquema dialéctico (no hegeliano) discutido en mi artículo "Qué es la Dialéctica", 1940, que ahora constituye el capítulo 15 de *Conjectures and Refutations*, 1963 [Traducción castellana de Néstor Míguez, *El Desarrollo del Conocimiento Científico. Conjeturas y Refutaciones*, Buenos Aires, Paidos, 1967, páginas 359-86].

[14] Utilizo aquí el término "transfondo" [*background*] en lugar de "conocimiento básico" [*background knowledge*], ya que no quiero entrar a discutir si es admisible o no un sentido objetivo (perteneciente al tercer mundo) del término *"conocimiento"*. (Véase, no obstante, *Conjectures and Refutations*, págs. 227 y sig. [trad. cit., pág. 264]. Para "conocimiento básico", véase *op cit.*, especialmente las págs. 112, 238 y si-

compuesto, al menos, por el *lenguaje* que siempre engloba diversas teorías en la misma estructura de sus usos (como ha señalado, por ejemplo, Benjamin Lee Whorf), junto con otras muchas suposiciones teóricas que, por el momento, no ponemos en tela de juicio. Los problemas sólo pueden surgir de un trasfondo semejante.

Un problema, juntamente con su trasfondo (y tal vez junto con otros objetos del tercer mudo), constituye lo que denomino *situación problemática*. Hay otros objetos del tercer mundo con los cuales operamos, tales como competición y conflicto (entre teorías, problemas, aspectos de las conjeturas, interpretaciones y posturas filosóficas), así como comparaciones, contrastes o analogías. Es importante señalar que la relación entre una solución y un problema es una relación lógica y, por tanto, una relación objetiva perteneciente al tercer mundo. También hay que señalar el hecho de que si nuestro intento de solución no resuelve nuestro problema, puede resolver otro problema que lo sustituya. Esto conduce a la relación perteneciente al tercer mundo denominada "*desplazamiento de problemas*" por I. Lakatos, que distingue entre desplazamientos de problemas progresivos y regresivos [15].

7. Comprensión y resolución de problemas

Quiero señalar ahora que la actividad de la comprensión es esencialmente la misma que la de cualquier resolución de problemas. Es indiscutible que, como toda actividad intelectual, consta de procesos subjetivos del segundo mundo. Con todo, el segundo mundo en cuestión puede y debe ser analizado en términos de una operación con elementos objetivos del tercer mundo. Dicha operación establece, en algunos casos, una especie de familiaridad con esos objetos y su manejo. Puede compararse —para emplear una analogía— con la actividad de un constructor de puentes o de casas: al intentar resolver un problema práctico, maneja u opera con unidades estructurales simples o complejas con ayuda de instrumentos simples o sofisticados.

Si reemplazamos estos instrumentos y unidades estructurales del primer mundo por instrumentos y unidades estructurales del tercero, como problemas, teorías o argumentos críticos, obtendremos una representación de lo que hacemos cuando tratamos de entender o aprehender una estructura

---

guientes [113 y 275 respectivamente]). El sentido objetivo de "conocimiento" se discute ampliamente en mi artículo "Epistemología sin Sujeto Cognoscente" (leído en Amsterdam en 1967) que ahora aparece reimpreso en el capítulo 3 de este volumen.

[15] Cf. I. Lakatos, "Changes in the Problem of Inductive Logic", en I. Lakatos (ed.), *The Problem of Inductive Logic*, 1968. Véase también ahora, I. Lakatos, "Criticism and the Methodology of Scientific Research Programmes", en I. Lakatos y A. Musgrave, *Criticim and the Growth of Knowlege*, 1970. [El artículo citado apareció en *Proc. Arist. Soc, 69*, págs. 149-86. Popper lo confunde con una versión más amplia titulada "Falsification and the Methodology of Scientific Research Programmes" aparecida en el colectivo Lakatos-Musgrave, citado por Popper; traducido ahora al castellano en Grijalbo, con prólogo de J. Muguerza (en preparación). N. T.]

del tercer mundo o cuando nos esforzamos por aportar al tercer mundo alguna que otra contribución consistente en la resolución de un problema. Pero obtenemos algo más que una simple representación. Mi tesis central es que todo análisis, intelectualmente significativo, de la actividad de comprender ha de proceder fundamentalmente, si no totalmente, mediante un análisis del manejo que hacemos de los instrumentos y unidades estructurales del tercer mundo.

Para hacer un poco más digerible esta tesis, tal vez se pueda recordar que estas unidades estructurales del tercer mundo son *inteligibles*, es decir, objetos posibles (o virtuales) de comprensión. No es, pues, de extrañar que si nos interesamos por los *procesos* de comprensión o algunos de sus resultados, tengamos que describir lo que hacemos o realizamos casi totalmente en términos de estos objetos de comprensión, los inteligibles y sus relaciones. Todo lo demás, como la descripción de nuestros sentimientos subjetivos de excitación, desánimo o satisfacción, puede ser muy interesante, pero carece de relevancia para nuestro problema; es decir, para la comprensión de los inteligibles, de los objetos o estructuras del tercer mundo.

No obstante, estoy dispuesto a admitir que hay ciertas actitudes o experiencias subjetivas que desempeñan un papel en el proceso de comprensión. Pienso en cosas tales como *centrarse* en algo: seleccionar el problema o teoría que estamos investigando o, por el contrario, *desestimar* una teoría, no tanto por que sea falsa cuanto por su carácter irrelevante o, digamos, irrelevante para la discusión en un cierto estadio, aunque en otro pueda ser interesante; e, incluso, tal vez, pasar por alto una teoría por ser falsa y demasiado irrelevante como para que la discutamos explícitamente. Considerado desde un punto de vista lógico, esto equivale a proponer que su falsedad e irrelevancia sean relegadas al "trasfondo" de la discusión.

Lo más normal es comunicar semejante propuesta de relegar una teoría o un problema (o una narración o un "proyecto") por medios expresivos o emocionales [16]. Se ve fácilmente que desde el punto de vista del manejo de los objetos del tercer mundo, estos medios funcionan como una especie de escritura taquigráfica: en principio podrían sustituirse por un *análisis* más detallado *de la situación problemática objetiva*. El problema es que este análisis puede ser complejo, puede llevar mucho tiempo y puede dar la impresión de que no vale la pena, ya que el problema consiste precisamente en establecer el hecho de que hay cosas irrelevantes.

Este bosquejo de análisis de algunas resonancias emocionales es un intento de ejemplificar la pretensión de que incluso esas resonancias puedan a menudo entenderse mejor en términos de objetos del tercer mundo, como son las situaciones problemáticas.

Esta pretensión no ha de confundirse con otra aún más importante,

---

[16] Se puede encontrar un buen análisis de esta situación en la crítica que hace Collingwood a Richards, ya mencionada anteriormente; véase *The Principles of Art*, 1938, especialmente las págs. 164 y sigs. En realidad, la crítica de Collingwood es un bello ejemplo de análisis del contenido emocional de un objeto del tercer mundo en términos de la situación problemática, su trasfondo y su solución.

según la cual la tarea de explicar estados psicológicos tales como las emociones, crea sus propios problemas teóricos que han de ser resueltos por sus propias teorías tentativas: teorías (es decir, objetos del tercer mundo) acerca del segundo mundo. Sin embargo, no hay que interpretar esto en el sentido de que sólo podamos entender a las personas estudiando teorías psicológicas acerca de ellas; ni tampoco pretende anular, ni siquiera restringir, mi tesis de que en toda comprensión, incluso en la de las personas y sus acciones y, por ende, *en la comprensión de la historia,* nuestra tarea suprema sea el análisis de las situaciones del tercer mundo.

Por el contrario, uno de los puntos más importantes de mi teoría es que las acciones y, por tanto, la historia, se pueden explicar como un proceso de resolución de problemas, por lo cual se les puede aplicar mi análisis en términos del esquema de *conjeturas y refutaciones* ($P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$, explicado más arriba en la sección 6).

Antes de proceder al desarrollo de este aspecto tan interesante, discutiré primero con algún detalle un proceso de comprensión de un objeto del tercer mundo: una simple ecuación aritmética.

8. Un ejemplo muy trivial

La igualdad entre 777 por 111 y 86.247 es un hecho aritmético muy trivial que se puede escribir en forma de ecuación. También se puede considerar como un teorema, así mismo trivial, de la teoría de los números naturales.

¿Acaso comprendemos esta proposición trivial?

Sí y no. Es evidente que comprendo la afirmación cuando la veo escrita, pues de otro modo no podría manejar o retener un número tan grande como es 86.247. (He realizado la experiencia y lo he confundido con el número 86.427.) Mas, en *cierto* sentido, la comprendo, evidentemente, tan pronto como la oigo: 777 y 111 son fáciles de manejar y comprendo que la proposición en cuestión suministra una *solución al problema*: ¿qué número es igual a 777 por 111 en el sistema en base diez?

Ya sé, por lo que respecta a la *resolución* del problema, que mucha gente es capaz de encontrar su solución mentalmente con toda facilidad, mientras que yo, para lograrlo, tengo que hacer un gran esfuerzo. Ahora bien, si lo que quiero es estar seguro del resultado obtenido de manera que no lo confunda al momento con otro distinto, he de echar mano de lo que Bridgman llama una "operación con lápiz y papel": he de ponerlo todo en un algoritmo en el que haya unidades estructurales fácilmente manejables. (Evidentemente, se trata de unidades estructurales del tercer mundo.) Una de las cosas que entran aquí es la *eliminación de errores*: las operaciones realizadas con papel y lápiz facilitan el diagnóstico y supresión de errores.

Hasta aquí, hemos utilizado tres de los cuatro objetos que aparecen en mi esquema de resolución de problemas (el esquema $P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$, introducido en la sección 6). Para comprender una proposición, una teoría tentativa, preguntamos en primer lugar, ¿De qué va el problema?, y para

eliminar cualquier error, calculamos con lápiz y papel. Si bien hemos partido de una teoría tentativa (*TT*), continuamos con el problema subyacente (P_1) y luego, con el método de cálculo diseñado para eliminar errores (*EE*). ¿Se presenta también un segundo problema (P_2)? Efectivamente: el método de eliminación de errores conduce a un desplazamiento de problemas. En nuestro caso se trata de un desplazamiento muy trivial y regresivo —la sustitución de un problema de multiplicar por tres multiplicaciones más sencillas y una suma. El desplazamiento (del problema P_1 al P_2) es, evidentemente regresivo, ya que, como es obvio, carece de todo interés teórico— se trata tan sólo de aplicar un procedimiento rutinario cuya única función es obtener la solución de una manera sencilla y fácil de comprobar (es decir, de eliminar errores).

Hasta en este ejemplo tan trivial podemos distinguir diversos grados de comprensión.

(1) La simple comprensión de lo dicho, en el mismo sentido en que "comprendemos" también la proposición "777 por 111 es igual a 68.427" sin constatar su falsedad.

(2) La comprensión de que constituye la solución de un problema.

(3) La comprensión del problema.

(4) La comprensión de que la solución es verdadera, cosa que en nuestro caso es trivialmente sencilla.

(5) La comprobación de la verdad del resultado por algún método de eliminación de errores, cosa que en nuestro caso es también trivial.

Está claro que hay otros grados de comprensión. En especial, la comprensión del problema (3) puede desarrollarse más, ya que solamente algunos comprenderán el carácter verbal del problema, en la medida en que "777 por 111" —aunque no esté escrito en base diez— es un medio tan excelente, o mejor, de construir un sinónimo del número "8 por 10.000 más 6 por 1.000 más 2 por 100 más 4 por 10 más 7", no siendo 86.247 más que un método abreviado de escribir este último número. Este tipo de comprensión es un ejemplo del intento de entender el *trasfondo* que normalmente se da por supuesto y, de este modo, *decubre un problema dentro de dicho trasfondo.*

Es evidente que, por regla general, estos grados de comprensión [17] no

[17] En muchas ocasiones señala Dilthey, con toda razón, la existencia de grados de comprensión. Sin embargo, no estoy muy seguro de que logre distinguir siempre los grados de comprensión (es decir, el carácter profundo o completo de la comprensión), de la certeza ("*Sicherheit*") de la comprensión, idea que considero, no sólo diferente, sino también equivocada. Así, por ejemplo, dice Dilthey: "El grado superior de certeza se alcanza en el campo de la interpretación de [los objetos del] espíritu científico". (W. Dilthey, *Gesammelte Schriften*, vol. 7, pág. 261.) Creo que hay una confusión, a menos que yo haya comprendido mal sus palabras. El hecho de que una *elevada certeza en la comprensión* pueda ir acompañada de un *grado de comprensión* extremadamente *reducido* es algo que podemos ver perfectamente si reflexionamos sobre la siguiente formulación de R. Carnap en su *Introduction to Semantics*, 1942, pág. 22: "...comprender una oración, conocer lo que afirma equivale a saber bajo qué condiciones sería verdadera". Realmente sé que la ecuación "777 por 111=86.427" será verdadera en el caso de que 777 por 111 sea de hecho igual a

pueden ponerse en un orden lineal, ya que en casi cualquier punto, sobre todo en los casos menos triviales, pueden surgir ramificaciones de nuevas posibilidades de mejor comprensión ulterior.

Así pues, con nuestro sencillo ejemplo, podemos aprender muchas cosas, la más importante de las cuales tal vez sea la siguiente. Siempre y cuando intentemos interpretar o comprender una teoría o proposición, por trivial que sea, como en el caso de la ecuación que discutíamos, de hecho estamos planteando un *problema de comprensión,* cosa que siempre termina convirtiéndose en *un problema* acerca de un problema; es decir, un *problema de nivel superior.*

9. Un caso de comprensión histórica objetiva [18]

Todo lo dicho se aplica a cualquier problema de comprensión y, en especial, al problema de la *comprensión histórica.* Mi tesis es que el objeto fundamental de toda comprensión histórica es la reconstrucción hipotética de una *situación problemática* histórica.

Intentaré explicar con algún detalle esta tesis, sirviéndome de otro ejemplo y con ayuda de unas pocas consideraciones históricas en torno a la *teoría de las mareas* de Galileo. Esta teoría ha resultado ser un "fracaso" (puesto que niega que la Luna tenga algún efecto sobre las mareas) e, incluso en su propia época, Galileo se vio atacado personalmente con rigor por su dogmatismo al aferrarse obstinadamente a una teoría tan claramente falsa.

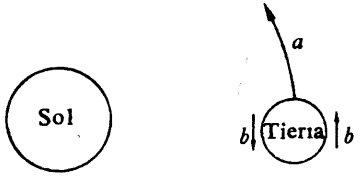
En resumen, la teoría de Galileo afirma que las mareas son un resultado de aceleraciones que, a su vez, resultan de los movimientos complejos de la Tierra. Concreamente, cuando a la rotación regular de la Tierra se suma el movimiento en torno al Sol, la velocidad de un punto de la superficie situado en un momento dado en la parte opuesta al Sol será mayor que la velocidad de ese punto cuando, al cabo de doce horas, esté de cara al Sol. (Ya que si a es la velocidad orbital de la Tierra y b la velocidad de rotación de un punto situado en el ecuador, entonces $a+b$ es la velocidad del punto a media noche y $a-b$, su velocidad a mediodía.)

En consecuencia, el cambio de velocidad significa que deben surgir aceleraciones y deceleraciones periódicas. Ahora bien, dice Galileo, toda

---

86.427 (cosa que no ocurre). Esto es algo que sé gracias a la definición de la verdad de Tarski y sé que esta especie de condición de verdad vale para *todo enunciado.* Por tanto, *he de comprender con certeza todo enunciado* si es que comprendo el lenguaje, lo cual es efectivamente verdad en el caso de un *grado de comprensión* extremadamente *bajo,* cosa que difícilmente puede ser lo que Carnap y Dilthey querían decir con su teoría.

[18] *En las restantes notas de este artículo trataré de poner de manifiesto, en relación con problemas de comprensión histórica, la superioridad del método del tercer mundo —consistente en construir críticamente las situaciones problemáticas— sobre el método del segundo mundo consistente en revivir intuitivamente una experiencia personal (método cuyo valor limitado y subjetivo no pretendo rechazar totalmente, ya que al mismo tiempo resulta inevitablemente sugestivo).*



deceleración y aceleración periódica de un recipiente con agua da lugar a fenómenos semejantes a las mareas. (Así formulada, la teoría de Galileo, aunque plausible, es incorrecta: aparte de la aceleración constante debida a la rotación terrestre —es decir, la aceleración centrípeta— que surge aunque *a* sea cero, no hay ninguna otra aceleración y, por tanto, concretamente, no hay ninguna aceleración periódica [19].)

¿Qué nos cabe hacer para mejorar nuestra comprensión histórica de esta teoría tan frecuentemente mal interpretada? La respuesta que ofrezco a este *problema de comprensión* (que designaré con "*Pc*") se desarrolla por vías semejantes a las de mi repuesta ya discutida en relación con nuestra ecuación aritmética trivial.

Propongo que el primer y más importante paso que hemos de dar es preguntarnos a nosotros mismos: *¿cuál era el problema (del tercer mundo)* a cuya solución provisional iba dirigida la teoría de Galileo? ¿Cuál era, además, la situación —la *situación problemática* lógica— en que se planteó el problema?

El problema de Galileo era muy sencillo: explicar las mareas. Sin embargo, la situación del problema era mucho menos simple.

Está claro que a Galileo no le interesaba de un modo inmediato lo que acabo de denominar su *problema*. Fue otro problema el que le condujo al de las mareas: el de la verdad o falsedad de la teoría copernicana, el de si la tierra está en reposo o en movimiento. Lo que Galileo pretendía era poder utilizar una teoría acerca de las mareas que tuviese éxito como argumento decisivo en pro de la teoría copernicana.

Lo que denomino *situación problemática* de Galileo resulta ser un asunto complicado. Está claro que la situación problemática entraña el problema de las mareas, otorgándole un papel específico: la explicación de las mareas ha de servir como piedra de toque de la teoría copernicana. Con todo, esta consideración no basta para comprender la situación problemática de

---

[19] Podría decirse que la teoría cinemática de las mareas de Galileo está en contradición con el llamado principio de relatividad de Galileo. Pero esta crítica sería falsa tanto histórica como teóricamente, ya que este principio *no* se refiere a los movimientos de *rotación*. La intuición física de Galileo —la rotación de la Tierra tiene consecuencias mecánicas no relativistas— era correcta; y aunque esas consecuencias (el movimiento de una peonza que rota, el péndulo de Foucault, etc.) no se apliquen a las mareas, al menos la fuerza de Coriolis influye algo en ellas. Además obtenemos aceleraciones cinemáticas (pequeñas) tan pronto como tenemos en cuenta la curvatura del movimiento de la Tierra en torno al Sol.

Galileo, puesto que su teoría tentativa no trataba solamente de explicar el cambio de mareas, sino que pretendía explicarlo situándolo contra determinado *trasfondo*, así como *en el seno de determinado marco teórico*. Aunque para Galileo el trasfondo no planteaba problemas, sí los planteaba, y en gran manera, lo que propongo denominar "el marco de Galileo", cosa de la cual Galileo era plenamente consciente.

Resulta así, que para resolver el *problema de comprensión* (P^c) hemos de investigar sobre un objeto del tercer mundo bastante complejo que consta del problema de las mareas (al que la teoría de Galileo ofrecía una solución tentativa) junto con su encuadre —su trasfondo y su marco: llamo *situación problemática* a este objeto complejo.

La *situación problemática* de Galileo puede caracterizarse del siguiente modo.

Como buen cosmólogo y teórico, Galileo se había visto atraído desde hacía tiempo por la increíble osadía y sencillez de la idea fundamental de Copérnico, según la cual tanto la Tierra como el resto de los planetas no son más que lunas del Sol. El poder explicativo de esta idea audaz era enorme, así que cuando Galileo descubrió las lunas de Júpiter, viendo en ellas un modelo reducido del sistema solar, interpretó su descubrimiento como una corroboración empírica de su audaz concepción, a pesar de su carácter tan especulativo y casi *a priori*. Además, había conseguido contrastar una predicción derivable de la teoría copernicana: predijo que los planetas más próximos al Sol que la Tierra tendrían fases como las de la Luna, cosa que Galileo logró observar en el caso de Venus.

Como la de Ptolomeo, la teoría de Copérnico era en esencia un modelo cosmológico de carácter geométrico, construido por medios geométricos (y cinemáticos). Pero Galileo era un físico. Sabía que el problema real era dar con una explicación física de índole mecánica (o tal vez transmecánica). De hecho descubrió algunos de los elementos de tal explicación, concretamente la ley de inercia y la ley correspondiente de conservación para los movimientos de rotación.

La audacia de Galileo le impulsó a intentar fundamentar su física en estas dos únicas leyes de conservación, a pesar de que era plenamente consciente de que en su conocimiento físico tenía que haber grandes brechas del tercer mundo. Desde el punto de vista del método, Galileo tenía pleno derecho a intentar explicarlo todo arrancando de este punto de partida tan estrecho, ya que sólo si intentamos explotar y contrastar hasta el límite nuestras teorías falibles podemos esperar aprender algo de su fracaso.

Esto explica por qué Galileo, a pesar de esar familiarizado con el trabajo de Kepler, se aferró a la hipótesis del movimiento circular de los planetas. Estaba en su derecho, teniendo en cuenta que este movimiento circular podía explicarse mediante sus leyes de conservación básica. Se dice a veces que intentó encubrir las dificultades de los círculos de Copérnico y que simplificó excesivamente la teoría copernicana de un modo injustificable, así como que debería haber aceptado las leyes de Kepler.

Pero todo esto revela un fallo de comprensión histórica —un error en el análisis de la situación problemática del tercer mundo—. Galileo tenía pleno derecho a operar con supersimplificaciones audaces, del mismo modo que las elipses de Kepler son supersimplificaciones igualmente audaces. Pero lo que pasa es que Kepler tuvo suerte por el hecho de que Newton utilizó, y explicó, en seguida su supersimplificación como contrastación de su solución al problema de los dos cuerpos.

¿Mas, por qué rechazó Galileo la idea de sobra conocida de la influencia de la Luna sobre las mareas? Esta pregunta revela un aspecto de suma importancia en la situación problemática. En primer lugar, Galileo rechazó la influencia lunar por su oposición a la astrología que identificaba esencialmente los planetas y los dioses. En este sentido, fue un pionero de la ilustración, así como un opositor de la astrología de Kepler, aunque le admirase [20] En segundo lugar, operaba con un principio de conservación mecánico para movimientos de rotación, lo que parecía excluir las influencias interplanetarias. Si Galileo no hubiese intentado explicar las mareas apoyándose en una base tan estrecha, nunca hubiésemos descubierto que lo era y que, por tanto, era necesario otra idea —la idea newtoniana de atracción (y con ella la de fuerza)—. Estas ideas eran muy semejantes a las ideas astrológicas que la mayoría de los ilustrados, como Berkeley [21], e incluso el propio Newton, consideraban ocultas.

De este modo, el análisis de la situación problemática de Galileo nos lleva a justificar la racionalidad de su método en muchos de los aspectos en que le han criticado diversos historiadores, con lo que accedemos a una mayor *comprensión histórica* de Galileo. Así pues, son superfluos los intentos que se han hecho de explicación psicológica, tales como la ambición, los celos, la agresividad o el deseo de provocar la agitación. Aquí los sustituimos por un análisis de la situación en el tercer mundo. De un modo semejante, es superfluo criticar el "dogmatismo" de Galileo por haberse aferrado al movimiento circular o introducir la idea de una atracción psicológica misteriosa en el "misterioso movimiento circular". (Dilthey dice de ella que es una idea arquetípica o que es psicológicamente atractiva [22].) Es superfluo, porque Galileo utiliza un método correcto cuando trata de

---

[20] Véase *Conjectures and Refutations*, pág. 188 [trad. cit., pág. 220].

[21] Véase *Conjectures and Refutations*, pág. 188 [220] y capítulo 6.

[22] Dilthey habla del "misterioso movimiento circular" ("*die geheimnisvolle Kreisbewegung*": *Schriften*, vol. I, págs. 95-6) de la antigua astronomía. Creo que se trata de una mala interpretación que habla en contra de sus *grados de certeza* discutidos en una nota anterior. (Tal vez Dilthey replicase que en este campo la ciencia no empieza hasta Newton y que, por tanto, se refería a ideas precientíficas. No creo que se pueda aceptar esta respuesta, negando a Galileo la condición de científico: la ciencia empieza con Anaximandro o incluso antes.)

Para un estudio breve, aunque profundo, de Ptolomeo *vs.* Copérnico, véase O. Neugebauer, *The Exact Sciences in Antiquity*, 1957, págs. 191 y sigs. (Por no haber conseguido separar tajantemente los problemas físicos de los geométricos, el propio Neugebauer condena —en la página 204— el empeño de Copérnico o Galileo de emplear círculos, considerándolos dogmáticos.)

llegar lo más lejos posible con la ayuda de la ley racional de conservación de los movimientos de rotación. (Aún no había una teoría dinámica.)

Este resultado puede ilustrar de qué manera aumenta nuestra comprensión histórica del papel de Galileo cuando comprendemos su situación problemática objetiva. Podemos referirnos con "P_1" a esta situación problemática, puesto que desempeña un papel análogo al P_1 del que hablábamos antes. Del mismo modo podemos llamar "*TT*" a la teoría tentativa de Galileo y "*EE*" a su intento, y al de tantos otros, de discutirla críticamente para eliminar sus errores. Galileo, aunque tenía esperanzas, no estaba ni mucho menos satisfecho con el resultado de su discusión. Podemos decir que su P_2 estaba muy próximo a su P_1; es decir, el problema seguía aún planteado.

Mucho más tarde, toda esta cuestión condujo a un cambio revolucionario (debido a Newton) en la *situación problemática* (P_2). Newton amplió el marco galileano —el marco de las leyes de conservación—, en cuyo seno se había concebido el problema de Galileo. En parte, el cambio revolucionario de Newton consistió en la readmisión de la Luna, cuya eliminación de la teoría de las mareas había sido una consecuencia del marco (y trasfondo) de Galileo.

Resumiré brevemente esta historia. El marco físico de Galileo era una versión un tanto simplificada del modelo copernicano para el sistema solar que estaba construido a base de un sistema de ciclos (y, tal vez, de epiciclos) dotados de una velocidad de rotación constante. El propio Einstein ha glosado su "adhesión a la idea de movimiento circular" achacándole la "responsabilidad del hecho de que no reconociese *plenamente* la ley de inercia y su fundamental significación" [23]. Con ello olvidaba que, del mismo modo que la teoría de Newton se basaba en la ley de inercia o en la de conservación del movimiento, la teoría de los ciclos y epiciclos en su versión simplificada partidaria de las velocidades constantes —versión preferida de Galileo— se fundaba originalmente en la ley de la conservación de momento angular. No cabe duda de que ambas leyes de conservación eran acepta-

[23] Las palabras citadas pertenecen a la página XI del admirable Prólogo de Einstein al *Diálogo sobre los dos Máximos sistemas del Mundo*, vertido [al inglés] por Stillman Drake, ed. revisada, 1962. Einstein reconoce que Galileo poseía la ley de inercia, aunque sin duda no reconoció *plenamente* (el subrayado es de Einstein) su significado fundamental. He de aludir aquí al hecho de que se ha criticado mucho a Galileo por presentar el sistema copernicano de una manera excesivamente simplificada, diciendo además con intención crítica que "Ptolomeo introduce grandes epiciclos" (*Op. cit.*, págs. 341 y sigs.) silenciando el hecho de que Copérnico los usa también. Es éste un problema de interpretación histórica. Me da la impresión de que Galileo dejó sin resolver conscientemente el problema planteado por el hecho de que el sistema copernicano simplificado, basado únicamente en el movimiento circular de velocidad constante sin epiciclos, no concordaba exactamente con las observaciones. Con todo, se sentía muy impresionado por el hecho de que concordase con ellos comparativamente muy bien y creía que los *problemas puramente geométricos* que quedaban planteados se resolverían *únicamente si se resolvían los problemas físicos.* (Sugiere que las posibles soluciones podrían realizarse mediante epiciclos *no tan "grandes"*, torbellinos o fuerzas magnéticas; cf. *op. cit.*, págs. 398 y sigs.) Esta idea resultó ser correcta y no hemos de olvidar que incluso la solución geométrica de Kepler no era más que una aproximación; es decir, una simplificación excesiva.

das "instintivamente" a causa, quizás, de algo así como la selección de las conjeturas bajo presión de la experiencia práctica, ya que, por lo que respecta a la ley del momento angular, puede haber sido decisiva la experiencia con ruedas de carro bien engrasadas. Hemos de recordar también que la vieja teoría de la rotación circular de los cielos (que surge de esta experiencia) acabó siendo sustituida por la conservación del momento angular terrestre, señal de que los ciclos no eran ingénuos ni misteriosos como aún hoy se piensa frecuentemente. En este marco —que hay que oponer al de los astrólogos— no puede haber interacción entre los cuerpos celestes, con lo que la teoría lunar de las mareas sostenida por los astrólogos había de ser rechazada por Galileo[24].

¿Podemos sacar algo en limpio de este ejemplo? Creo que sí.

Primero, el ejemplo muestra la inmesa importancia de reconstruir la situación problemática (P_1) de Galileo para comprender su teoría (*TT*). La reconstrucción que hacemos para comprender las teorías que han fracasado, como la de Galileo, es aún más importante que las que hacemos para comprender las que han tenido éxito, ya que su fallo (el fallo de *TT*) puede explicarse por su fallo en el marco o en el trasfondo de P_1.

Segundo, en este caso es evidente que la reconstrucción de la situación problemática (P_1) de Galileo posee a su vez el carácter de una *conjetura* (así como de una simplificación o idealización excesivas): esto es obvio si tenemos en cuenta que mi análisis de esta situación problemática (P_1), breve como es, se aleja considerablemente del de otras personas que han tratado de comprender esta teoría fracasada de Galileo. Ahora bien, si mi reconstrucción de P_1 es una conjetura, *¿cuál es el problema que intenta resolver?* Evidentemente, se trata de P^c, el *problema de comprender* la teoría de Galileo.

He aquí mi *tercera* consideración: *nuestro problema de comprensión,* P^c, está *a un nivel superior* al de P_1. Es decir, el problema de la comprensión es un *metaproblema*: versa acerca de *TT*, así como acerca de P_1. De acuerdo con ello, la teoría ingeniada para resolver el problema de la comprensión será una *metateoría,* puesto que su misión consiste, en parte, en descubrir, para cada caso concreto, cómo eran realmente P_1, *TT*, *EE* y P_2.

Dicho sea de paso, no se ha de suponer que esto entrañe que en todos los casos particulares la metateoría *sólo* haya de descubrir las estructuras de P_1, *TT*, etc., aceptando sin discusión el esquema mismo ($P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$). Por el contrario, habría que subrayar una vez más que dicho esquema

---

[24] La conexión entre la teoría de las mareas de Galileo y su rechazo de la astrología se discute e interpreta en mi libro *Conjectures and Refutations,* nota 4 del capítulo I (pág. 38 [trad. cit., pág. 49]), así como en la nota 4 del capítulo 8 (pág. 188 [pág. 220]). Es una interpretación conjetural típica (en el sentido de mi *Open Society and its Enemies,* 1945, vol. I, capítulo 10, pág. 117 [trad. cit., pág. 185]) que como tal puede "iluminar el material histórico": me ha ayudado *a comprender mejor* el último pasaje del *Diálogo* de Galileo (op. cit., pág. 462; para la actitud de Galileo hacia la astrología véanse también las págs. 109 y sigs.) en el que alude a Kepler reprochándole sus "puerilidades" astrológicas.

es una simplificación excesiva que ha de ser elaborada e incluso radicalmente modificada cuando sea necesario.

El *cuarto* punto afirma que todo intento (excepto los más triviales) de comprender una teoría se ve abocado a abrir una investigación histórica en torno a dicha teoría y su problema que, de este modo, pasa a formar parte del objeto de la investigación. Si se trata de una teoría científica, la investigación versará sobre la historia de la ciencia. Si es histórica, pongamos por caso, la investigación versará sobre la historia de la historiografía. Los problemas a cuya solución se dirigen estas investigaciones históricas son metaproblemas que han de ser tajantemente separados de los problemas que constituyen el objeto de la investigación.

El *quinto* punto señala que la historia de la ciencia *no* ha de tomarse como una historia de las teorías, sino como una historia de las situaciones problemáticas y sus modificaciones (ora imperceptibles, ora revolucionarias) a causa de los intentos de resolver los problemas. Históricamente, los intentos que no se ven coronados por el éxito pueden resultar, así pues, tan importante como los intentos fructuosos en orden al desarrollo ulterior.

Según el *sexto* punto (que no hace más que desarrollar el tercero) hemos de distinguir claramente los metaproblemas y metateorías del historiador de la ciencia (que se sitúan al nivel P^c) de los problemas y teorías de los científicos (que se sitúan al nivel P_1). Es muy fácil confundirlos, porque si formulamos el problema del historiador diciendo: "¿Cuál era el problema de Galileo?", parece que la repuesta debería ser "P_1"; pero P_1 (en contraste con "el problema de Galileo era P_1") parece corresponder al nivel de objeto más bien que al meta-nivel [25], con lo que ambos niveles se confunden.

Ahora bien, en general, no hay ningún problema que pertenezca a ambos niveles, cosa que se puede ver con toda facilidad, ya que dos metateorías acerca del mismo objeto son a menudo muy distintas. Dos historiadores de la ciencia, aunque estén de acuerdo en "los hechos", pueden entenderlos o interpretarlos de modos muy distintos (que alguna vez se complementarán y otras estarán en conflicto). Incluso pueden estar en desacuerdo acerca de qué es lo que constituye su problemática. Así pues, en general, no compartirán los mismos problemas ni, con más razón, la misma teoría objeto de su investigación e interpretación.

Además, a la hora de interpretar una teoría, el metateórico es libre de utilizar todo lo que pueda ayudarle. Así, por ejemplo, puede comparar la teoría con otras teorías rivales radicalmente distintas. *Por tanto, algunas de las unidades estructurales del tercer mundo que constituyen la teoría pueden ser manifiestamente distintas de las que constituye la teoría a interpretar o comprender.*

Esta cuestión es importante porque deja bien sentado *a fortiori* que aunque podamos hablar razonablemente (cosa que me atrevería a negar)

---

[25] De hecho, la respuesta es una conjetura histórica *acerca del* problema de Galileo (P_1). Más adelante discutiremos con mayor detalle los metaproblemas del historiador y sus respuestas conjeturales.

de algo así como una *semejanza* entre los *contenidos de pensamiento* del tercer mundo por un lado, y los *procesos de pensamiento* del segundo mundo mediante los que captamos aquellos contenidos, por otro, aún entonces se podría rechazar en general la existencia efectiva de una semejanza —a ningún nivel de problemas— entre los contenidos y los procesos de pensamiento correspondientes. Esto es así, porque el método (del tercer mundo) de comprensión histórica que intento describir es tal que, de ser posible, sustituye las explicaciones psicológicas por el análisis de las *relaciones del tercer mundo*: en lugar de emplear principios explicativos de carácter psicológico, utilizamos consideraciones propias del tercer mundo de carácter fundamentalmente lógico. Según mi tesis, nuestra comprensión histórica puede desarrollarse gracias a tales análisis.

El *séptimo* punto, que tal vez sea el más importante, se refiere a lo que algunas veces he denominado *lógica de la situación o análisis de la situación* [26]. (El segundo nombre puede ser preferible, ya que el primero parece sugerir una teoría determinista de la acción humana que no corresponde en absoluto a mi intención.)

Por análisis de una situación entiendo una especie de explicación tentativa o conjetural de una acción humana que alude a la situación en que se encuentra el agente mismo. Puede ser una explicación histórica: tal vez podamos explicar cómo y por qué se creó una determinada estructura de ideas. Es evidente que nunca se podrá explicar plenamente una acción creadora. No obstante, podemos intentar obtener mediante conjeturas una reconstrucción ideal de la *situación problemática* en que se encontraba el agente, haciendo "comprensible" (o "racionalmente comprensible") la acción en ese sentido; es decir, haciendo una reconstrucción *adecuada a su situación tal como él la veía*. Este método de análisis situacional puede considerarse una aplicación del *principio de racionalidad*.

Una de las tareas del análisis de situaciones sería establecer la distinción entre la situación tal como la ve el agente y la situación tal como es (ambas, evidentemente, conjeturadas) [27]. Por tanto, el historiador de la

---

[26] He descrito el método de la *lógica de la situación* o del *análisis de la situación* en mi libro *The Open Society*, vol. II, capítulo 14, especialmente la pág. 97 [trad. cit., pág. 298], así como en *Poverty of Historicism*, 1057, secciones 31 ("La lógica de la situación en historia"; especialmente la pág. 149 [trad. cit., pág. 180]) y 32.

[27] En muchos casos podemos reconstruir, *objetivamente* (aunque de modo conjetural), (a) la *situación tal como era* y (b) una *situación* muy distinta *tal como la veía, comprendía o interpretaba* el agente. Es interesante que esto sea posible *incluso en la historia de la ciencia*. La mecánica ondulatoria de Schrödinger suministra un buen ejemplo, ya que éste no interpretó estadísticamente su problema (cosa que se vio claramente una vez que Born formuló su famosa "*interpretación estadística*"; véase mi artículo "Quantum Mechanics without The Observer' ", en Mario Bunge (ed.), *Quantum Theory and Reality*, 1967, que ahora forma parte de mi *Philosophy and Physics*, cap. 3). Pero, hay muchos otros ejemplos antiguos y recientes. Kepler entendía su problema como el descubrimiento de una pitagórica *armonía del mundo*. Einstein formuló el problema de la relatividad general sirviéndose de un postulado de covarianza; y, aunque admitió la crítica de E. Kretschmann (*Ann. Physik, 35*, pág. 575, 1917) quien decía que su postulado era vacío, creía evidentemente que podría reformularse de modo que sirviese para el fin propuesto, si bien nunca nos suministró una reformu-

ciencia, no sólo intenta explicar mediante el análisis de las situaciones, la teoría adecuada que propone un científico, sino que incluso trata de explicar su fracaso.

En otras palabras, nuestro esquema de resolución de problemas mediante conjetura y refutación, u otro similar, puede utilizarse como teoría explicativa de las acciones humanas, ya que podemos interpretar una acción como un intento de resolver un problema. Así, la teoría explicativa de la acción consistirá fundamentalmente en una reconstrucción conjeturada del problema y su trasfondo. Una teoría de este tipo puede ser perfectamente contrastable.

He intentado responder a la pregunta: "*¿Cómo podemos comprender una teoría o mejorar la comprensión que tengamos de ella?*", sugiriendo que mi repuesta en términos de problemas y situaciones problemáticas puede aplicarse a campos muy diversos del de las teorías científicas. En algunos casos, por lo menos, podemos aplicarla a las obras de arte: podemos hacer conjeturas acerca de cuál es el problema del artista y hemos de ser capaces de apoyar esta conjetura en elementos de juicio independientes. Este análisis puede ayudarnos a comprender la obra [28].

(La tarea de reconstruir una obra de arte estropeada —por ejemplo, un poema descubierto en un papiro deteriorado— puede considerarse quizás como algo intermedio entre la tarea de interpretar una teoría científica y la de interpretar una obra de arte.)

10. El valor de los problemas

Puede objetarse que la solución que sugiero al problema de "¿Cómo podemos comprender una *teoría* científica o mejorar la comprensión que tengamos de ella?", no hace más que aplazar la cuestión, puesto que lo único que consigue es sustituirla por otra semejante, como es: "¿Cómo podemos comprender un *problema* científico o mejorar la comprensión que de él tengamos?".

---

lación satisfactoria. En *Conjectures and Refutations*, capítulo 2, sección X, especialmente págs. 94-6 [trad. cit., págs. 112-4], analizo un ejemplo sacado de la filosofía (relacionado con el problema kantiano. "¿Cómo es posible la ciencia natural pura?").

[28] Se pueden encontrar muchos análisis de este tipo en las obras de E. H. Gombrich. Su libro *Art and Illusión*, 1959, es en parte (no totalmente) un estudio del impacto que sobre el Arte Occidental tuvieron los problemas planteados por la pretensión, alimentada en el pasado por muchos artistas, de crear una ilusión de realidad (utilizando, por ejemplo, la perspectiva). En su libro *Norm and Form*, 1966, pág. 7, cita la propia descripción que hacía Ghiberti de sus objetivos: "Me esforzaba... en imitar la naturaleza todo lo posible con todas las líneas que se encuentran en ella... [los paneles] no son más que entramados que el ojo calibra de modo tan verdadero, que, a cierta distancia, parecen tener relieve". Comenta Gombrich: "El artista trabaja como un científico, sus obras no sólo existen por cuenta propia, sino que además demuestran la solución de ciertos problemas". Evidentemente, esto forma parte del análisis de la obra de un artista y el hecho de que sobre otro artista puedan hacerse comentarios similares no indica que sus problemas sean semejantes. Por el contrario, los problemas cambian: las soluciones a los viejos problemas —el viejo problema de

La objeción es válida, aunque, por regla general, el desplazamiento del problema será progresivo (para utilizar la terminología del profesor Lakatos). Por regla general, la segunda pregunta —el metaproblema de comprender un problema— será más difícil y más interesante que la primera. En todo caso, considero que es la más fundamental de las dos, pues pienso que la ciencia parte de problemas (más bien que de observaciones e incluso teorías, aunque es obvio que el "trasfondo" del problema contendrá teorías y mitos).

Sea como sea, digo que este segundo metaproblema es distinto del primero. Evidentemente, siempre podremos y deberemos atacarlo como al primero —mediante una construcción histórica ideal— si bien creo que esto es insuficiente.

Mi tesis es que para obtener una *comprensión real* de un problema dado (la situación problemática de Galileo, por ejemplo) se precisa algo más que un análisis de este problema o de cualquier otro para el cual dispongamos de una buena solución: para comprender ese problema "muerto" tenemos que haber luchado, al menos alguna vez en la vida, con un problema vivo.

Por tanto, mi solución al metaproblema "¿Cómo podemos aprender a comprender un problema científico?" es la siguiente: aprendiendo a comprender algún problema *vivo*, lo cual, según sostengo, sólo puede hacerse *intentando resolverlo y fracasando en ello*.

Supongamos que un joven científico se encara con un problema que no comprende. ¿Qué es lo que puede hacer? Sostengo que aunque no lo comprenda, puede intentar resolverlo *criticando él mismo su solución* (o consiguiendo que la critiquen otros). Puesto que no entiende el problema, su solución fallará, cosa que la crítica se encargará de poner de manifiesto. De esta manera se dará un primer paso hacia la *localización de la dificultad*, lo cual indica, precisamente, que se dará un primer paso hacia la comprensión del problema. La razón de ello es que un problema constituye una dificultad, con lo que comprender un problema consiste en descubrir que hay una dificultad y dónde la hay, lo cual se consigue solamente averiguando por qué *no funcionan ciertas soluciones aparentes*.

Así pues, aprendemos a entender un problema intentando resolverlo y fracasando y, cuando hayamos fracasado cien veces, podremos convertirnos en unos expertos en ese problema particular. Es decir, si alguien propone una solución, podremos ver en seguida si su propuesta tiene futuro o si

---

crear ilusión de realidad o "naturaleza", por ejemplo— pueden provocar el rechazo del viejo problema y la búsqueda de otros nuevos.

Uno de estos nuevos problemas es cómo interesar al espectador y comprometerle a cooperar activamente, proponiéndole, por ejemplo, problemas de interpretación o reconstrucción. Cf. E. H. Gombrich, *Meditations on a Hobby Horse*, 1963 [*Meditaciones sobre un Caballo de Juguete*, Barcelona. Seix-Barral, 1968].

He de decir aquí que los análisis de Gombrich arrojan luz sobre el problema de lo que podríamos denominar *la autonomía de la obra de arte*: el hecho de que, aunque sea producida por el hombre, crea sus propias interrelaciones. (Véase también mi *Sobre Nubes y Relojes*, sección 24 y nota 65). Según una hermosa anécdota, cuando Haydn escuchaba el primer coro de su *Creación*, rompió en llanto mientras exclamaba: "Yo no he compuesto eso".

fallará a causa de esas dificultades que tan bien conocemos por nuestros pasados fracasos.

Así pues, aprender a entender un problema es cuestión de manejar unidades estructurales del tercer mundo, y obtener una aprehensión intuitiva del problema no es más que familiarizarse con esas unidades y sus interrelaciones lógicas. (Naturalmente, todo esto es semejante a la obtención de una aprehensión intelectual de una teoría.)

Sostengo que sólo quien haya luchado a brazo partido con un problema vivo podrá lograr una buena comprensión de un problema como el de Galileo; pues sólo él podrá evaluar su propia comprensión. Además, solamente él comprenderá plenamente (en el tercer nivel, como si dijéramos) el significado de mi afirmación de que el primer paso vital hacia la comprensión de una teoría es comprender la situación problemática a partir de la cual surge.

También sostengo que el tan debatido problema de la transferencia del aprendizaje de una disciplina a otra está íntimamente relacionado con el aumento de experiencia en la lucha con problemas vivos. A aquellos que sólo han aprendido a aplicar un marco teórico dado a la resolución de problemas que surgen dentro del ámbito de dicho marco, siendo solubles en él [29], es de esperar que su entrenamiento no les sea de mucha utilidad en otra especialidad. La situación es distinta en el caso de quienes han luchado con los problemas personalmente, en especial si su comprensión, clarificación y formulación se ha mostrado difícil [30].

Por tanto, pienso que quienes han luchado con un problema se pueden ver compensados obteniendo la comprensión de dominios alejados del suyo.

Puede ser interesante y fructífero investigar hasta qué extremos podemos aplicar el análisis situacional (la idea de resolución de problemas) al arte, la música y la poesía y si ello nos puede ayudar a comprender estos campos. No me cabe duda de que en muchos casos nos puede ayudar. Los cuadernos de Beethoven acerca del último movimiento de la Novena Sinfonía muestran que la introducción al movimiento contiene la historia de su intento por resolver un problema —el problema de prorrumpir en palabras—. Ver esto nos ayuda a comprender la música y al músico. El que esta comprensión nos ayude a gustar de la música es una cuestión diferente.

---

[29] Los científicos aquí descritos son los que practican lo que Thomas Kuhn ha llamado "ciencia normal", *The Structure of Scientific Revolutions*, 162 (2.ª edición, 1971) [traducción castellana de Agustín Contín, *La Estructura de las Revoluciones Científicas*, México, Fondo de Cultura Económica, 1971].

[30] En las dos o tres primeras páginas del capítulo 2 de *Conjectures and Refutations* he intentado defender la idea de que *no hay temas, sino tan sólo problemas* los cuales, evidentemente, aunque pueden provocar el surgimiento de teorías, *casi siempre exigen para su solución la colaboración de teorías muy diversas*, lo que muestra el carácter frustrante de la especialización.

11. La comprensión ("hermenéutica") en las humanidades

Esta última consideración me lleva a plantear el problema de la comprensión en las humanidades (*Geisteswissenschaften*).

Casi todos los grandes estudiosos de este problema —sólo mencionaré a Dilthey y Collingwood— sostienen que las humanidades difieren radicalmente de las ciencias naturales, siendo la diferencia más notoria el que las humanidades tienen como tarea fundamental el *comprender*, en un sentido según el cual podemos comprender a los hombres y no a la naturaleza.

Se dice que la comprensión se basa en nuestro carácter humano común. Su rasgo característico viene dado por una especie de identificación intuitiva con los demás hombres, mediante la ayuda de movimientos expresivos como son los gestos y el habla. Es además una comprensión de acciones humanas y, en último término, de los productos de la mente humana.

Hemos de admitir que, en el sentido aquí apuntado, podemos comprender a los hombres junto con sus actos y productos, si bien no nos es dado comprender la "naturaleza" —sistemas solares, moléculas o partículas elementales—. Con todo, no hay aquí una distinción tajante. Podemos aprender a comprender los movimientos expresivos de los animales superiores en un sentido muy semejante a aquel según el cual comprendemos a los hombres. Pero, ¿qué es un animal "superior"? ¿Acaso se limita a ellos nuestra comprensión? (H. S. Jennings aprendió a comprender lo suficiente bien los organismos unicelulares como para atribuirles propósitos e intenciones [21].) En el otro extremo de la escala, incluso la comprensión intuitiva de nuestros amigos no es perfecta ni mucho menos.

Estoy totalmente dipuesto a aceptar la tesis de que el objeto de las humanidades es la comprensión. Pero no veo claro que se pueda negar que lo sea también de las ciencias naturales. Evidentemente, se tratará de una "comprensión" en un sentido ligeramente distinto. Pero ya hay muchas diferencias en la comprensión de los hombres y sus acciones y no debemos olvidar afirmaciones como la siguiente, hecha por Einstein en una carta a Born:

"Usted cree en un Dios que juega a los dados y yo, en la absoluta regulación que introduce la ley en un mundo compuesto por una realidad objetiva y que intento captar de un modo salvajemente especulativo" [22].

Estoy seguro que los intentos salvajemente especulativos que hacía Einstein por "captar" la realidad son intentos de *comprenderla*, en un sentido tal de la palabra, que tiene por lo menos *cuatro* puntos de contacto con la comprensión de las humanidades. (1) Del mismo modo que comprendemos a otras personas por nuestra humanidad común, podemos comprender la naturaleza ya que formamos parte de ella. (2) Del mismo modo que comprendemos a los hombres en virtud de la racionalidad de sus pensamientos y

---

[21] Cf. H. S. Jennings, *The Behaviour of the Lower Organisms*, 1906.

[22] La carta de Einstein aparece citada (en el alemán original y en versión inglesa) en Max Born, *Natural Philosophy of Cause and Chance*, 1949, pág. 122.

acciones, podemos comprender las leyes de la naturaleza merced a cierto tipo de racionalidad o necesidad comprensible [33] inherente a ellas. Ha sido éste el propósito explícito de la mayoría de los grandes científicos desde Anaximandro al menos, para no hablar de Hesíodo o de Herodoto [34]; deseo que se ha visto satisfecho, siquiera sea momentáneamente, primero con la teoría gravitatoria de Newton y luego con la de Einstein. (3) La alusión a Dios en la carta de Einstein apunta a otro sentido compartido con las humanidades —el intento por comprender el mundo de la naturaleza al modo en que comprendemos una obra de arte: como una creación. Y (4) hay en las ciencias naturales esa conciencia de fracaso en última instancia de todos nuestros esfuerzos por comprender que tanto han discutido los humanistas atribuyéndolo a la "alteridad" de las demás personas, a la imposibilidad de la comprensión real y a la inevitable simplificación inherente a todo intento de comprender algo único y real. (Podemos añadir aquí que no importa mucho que esa realidad sea cósmica o microcósmica.)

Me opongo, por tanto, al intento de proclamar que el método de comprensión sea característico de las humanidades, sea la señal que sirve para distinguirla de las ciencias naturales. Y si sus partidarios denuncian mi punto de vista tildándolo de "positivista" o "cientifista" [35], tal vez pueda responder que ellos mismos parecen aceptar, *implícita y acríticamente*, que el positivismo o cientifismo es *la única filosofía adecuada a las ciencias naturales*.

Es algo comprensible, teniendo en cuenta que tantos científicos naturales han abrazado esta filosofía cientifista. Sin embargo, los humanistas deberían haber tenido una visión más penetrante de las cosas. Después de todo, la ciencia no es más que una rama de la literatura y trabajar en ciencia es una actividad humana como la construcción de una catedral. No

---

[33] Podemos mencionar exigencias de racionalidad tales como los principios de simetría (que han puesto especialmente de relieve Hermann Weyl y E. P. Winger), así como también ciertas ideas del tipo del "principio de acción y reacción de Eisntein" (podría llamársele también su "principio de realidad"), según el cual, el espacio y tiempo newtonianos son insatisfactorios porque pueden ejercer un efecto en los cuerpos, no estando a su vez sujetos a ningún tipo de efecto complementario (como ocurre con los campos).

[34] La idea de simetría cósmica puede encontrarse en la *Teogonía* de Hesíodo (720-5), en la teoría de Anaximandro sobre el tamaño y posición de la Tierra y en el intento de Herodoto por introducir algún tipo de simetría en una geografía que se mostraba, en gran medida, asimétrica. (Hizo que el Danubio y el Nilo fuesen lo más simétricos posible.) Cf. la nota anterior. Además, todo intento de introducir una medida de justicia, premio o castigo en el universo (Anaximandro, Herodoto) no es sino un esfuerzo por encontrar en él alguna racionalidad para poder comprenderlo.

[35] El término "cientifismo" significaba en un principio "la imitación servil del método y lenguaje de la ciencia [natural]", fundamentalmente por parte de los científicos sociales; lo introdujo en este sentido Hayek en su artículo "Scientism and the Study of Society", recogido ahora en su libro *Counter-Revolution in Science*, 1962. En *The Poverty of Historicism*, pág. 105. [trad. cit., pág. 98] sugería utilizarlo para designar la imitación totalmente *equivocada* del método de la ciencia. Actualmente, Hayek está de acuerdo (en el Prefacio a su libro *Studies in Philosophy, Politics, and Economes* que contiene un agradecimiento muy generoso) en que los métodos efectivamente utilizados por los científicos naturales son distintos de "lo que la mayoría nos dicen... a los representantes de otras disciplinas, instándonos a imitarlos".

cabe duda de que en la ciencia natural comtemporánea hay demasiada especialización y profesionalización que la convierte en algo inhumano; pero, desgraciadamente, se puede decir lo mismo y casi en la misma medida, de la psicología o la historia contemporáneas.

Además, hay un importante campo histórico —quizás el más importante— compuesto por la historia de la opinión y del conocimieneto humano que engloba la historia de la religión, de la filosofía y de la ciencia. Ahora bien, hay dos cuestiones acerca de la historia de la ciencia. Una de ellas es que sólo quien comprende la ciencia (es decir, los problemas científicos) puede comprender su historia; la otra, que sólo quien posee una comprensión real de su historia (la historia de sus situaciones problemáticas) puede comprender la ciencia.

La elaboración de las diferencias entre ciencia y humanidades ha sido durante mucho tiempo una moda que ha terminado por convertirse en una pesadez. Ambas practican el método de resolución de problemas, el método de conjeturas y refutaciones que es utilizado tanto para reconstruir un texto deteriorado como para construir una teoría acerca de la radioactividad [36].

Pero podría llegar incluso a acusar, al menos a algunos profesionales, de "cientifismo"; es decir, de intentar copiar los métodos de la ciencia natural, *no tal como es realmente*, sino tal como ellos creen equivocadamente que es. Este pretendido aunque inexistente método consiste en recolectar observaciones para "sacar conclusiones" a partir de ellas. Algunos historiadores lo imitan servilmente, pues creen que pueden recoger elementos de juicio documentales que constituirían, de modo similar a las observaciones de la ciencia natural, la "base empírica" de sus conclusiones.

Tal supuesto método nunca se puede poner en práctica: no se puede ni reunir observaciones ni elementos de juicio documentales si no se parte de un problema. (Un coleccionista de billetes reúne documentos, aunque rara vez reúne elementos de juicio histórico.)

Peor aún que el intento de aplicar un método inaplicable es la idola-

---

[36] Evidentemente, hay muchísimas diferencias. Pero, pocas cosas hay tan similares a determinados procedimientos de la física teórica como la reconstrucción conjetural de un texto deteriorado. Una conjetura de este tipo es siempre contrastable y algunas de ellas han sido refutadas. (Véase, por ejemplo, el *Papiro de Berlín*, número 9777 que J. U. Powell puso en relación con el *Papiro de Oxyrhynco*, XVII, 2075, fr. 1, para poder refutar ciertas reconstrucciones conjeturadas.) No obstante, estos casos son más bien raros. Por regla general, "Las contrastaciones de... [la mayoría de] las interpretaciones históricas" (como las que se pueden encontrar en el libro de J. W. N. Watkins, *Hobbes Systen of Ideas*, 1965, o en las págs. 248-53 y 319 del volumen I, así como en otros lugares de mi *Open Society*) "no pueden ser nunca tan rigurosas como las de una... [hipótesis] física", como decía en la pág. 171 [trad. cit., pág. 185]. Debería haber señalado que las hipótesis más interesantes —las cosmológicas— constituyen una excepción. Evidentemente, algunas de ellas pueden ser contrastadas, incluso con la precisión requerida para refutarlas; pero otras, y por cierto muy interesantes, parecen ser incontrastables por siempre. (Sobre la contrastación véase mi libro *Logic of Scientific Discovery*, 1969, cuya primera edición se publicó con el título *Logik der Forschung*, 1934. [Versión castellana de V. S. de Zavala. *La Lógica de la Investigación Científica*, Madrid. Tecnos, 1962].)

tría de una especie de conocimiento infalible y digno de confianza que estos historiadores confunden con el ideal de la ciencia [37]. Es evidente que todos nosotros ponemos todos los medios a nuestro alcance para evitar el error y que debemos entristecernos cuando nos equivocamos. Sin embargo, el ideal de evitar el error es un ideal pobre: si no tenemos la osadía de abordar problemas difíciles en los que el error sea casi inevitable, el conocimiento no aumentará. De hecho, son las teorías más audaces, *incluso las erróneas,* las que nos enseñan. Nadie está libre de cometer errores; lo grande es aprender de ellos [38].

12. Comparación con el método de la evocación subjetiva de Collingwood

A fin de ilustrar la aplicación del análisis de situaciones a la historia y para compararlo con el método de comprensión subjetiva (perteneciente al segundo mundo), empezaré citando un pasaje de R. G. Collingwood, el filósofo, historiador e historiógrafo.

Citaré este pasaje de Collingwood, porque puedo acompañarlo un largo trecho, aunque no hasta el final. Nos separamos en lo que atañe al segundo y tercer mundo: la elección de un método objetivo o subjetivo. (Estamos de acuerdo acerca de la significación de las situaciones problemáticas.) La manera subjetiva que tiene Collingwood de plantear las cosas no se reduce a un simple problema de formulación, sino que por el contrario constituye un aspecto esencial de su teoría de la comprensión (como ocurre con Dilthey, si bien éste intentó desembarazarse de la subjetividad por temor a la arbitrariedad) [39].

Como pone de manifiesto el pasaje, la tesis de Collingwood es que la comprensión que el historiador tiene de la historia consiste en su evocación de experiencias pasadas:

> Supongamos... que [el historiador] está leyendo el Código de Teodosio y que tiene ante sí determinado edicto de un emperador. El simple hecho de leer el escrito y

---

[37] Cf. mi libro *Logic of Scientific Discovery,* sección 85.

[38] Este es el tema principal de mi libro *Conjectures and Refutations;* véase el *Prefacio.*

[39] Se trata de uno de los mayores problemas de Dilthey. Hablaba especialmente de la necesidad de superar las tendencias subjetivistas y escépticas en historiografía. A este respecto, hay que aludir al famoso problema que Dilthey y otros denominaban "el círculo hermenéutico": el problema de que el todo (un texto, un libro, una obra filosófica, un período) sólo puede comprenderse si comprendemos las partes constituyentes, cuando estas partes sólo pueden ser comprendidas, a su vez, si comprendemos el todo. Normalmente, se ignora que Bacon formuló esta idea con toda claridad en *De Argumentis,* VI, x. VI. [trad. castellana, *Del Adelanto y Progreso de la Ciencia,* Buenos Aires, Lautaro, 1947, VI, XLVI]: "Del conjunto de las palabras hemos de extraer el sentido a cuya luz ha de interpretarse cada palabra particular". ("Interpretar" significa aquí simplemente "leer"; véase la última nota.) También se encuentra esta misma idea expuesta irónicamente en el *Diálogo* de Galileo (op. cit., pág. 108) en el que el autor hace decir a Simplicio que para comprender a Aristóteles hay que tener "siempre presente en la mente todo lo que ha dicho".

poder traducirlo no equivale a conocer su alcance histórico. Para ello deberá enfrentarse a la *situación* que el emperador trataba de controlar del mismo modo que se enfrentaba a ella el emperador. A continuación, debe ver por su cuenta, como si la *situación* del emperador fuese la suya propia, de qué manera se podría dominar dicha *situación;* ha de ver las alternativas posibles y las razones para elegir una de ellas más bien que otra; es más, debe pasar por los mismos procesos por los que pasó el emperador al decidirse por esta alternativa particular. Por tanto, evoca en su propia mente la experiencia del emperador, y sólo en la medida en que logre hacerlo poseerá un conocimiento histórico del sentido del edicto que sea algo más que un mero conocimiento filológico [40].

Como se ve, Collingwood pone el acento en la *situación* de un modo muy semejante a como yo hago hincapié en lo que denomino *situación problemática.* Con todo, hay una diferencia. Collingwood deja bien claro que lo esencial al comprender la historia no es el análisis de la situación misma, sino el proceso que tiene lugar en la mente del historiador y que consiste en la evocación, en la repetición simpatética de la experiencia original. Para Collingwood, el análisis de la situación no es más que una ayuda —aunque indispensable— para esta evocación. Mi punto de vista es diametralmente opuesto. Considero que el proceso de evocación psicológica no es esencial, si bien admito que algunas veces puede ser de utilidad para el historiador, en cuanto comprobación intuitiva del éxito de su análisis de la situación. *Lo que considero esencial no es la evocación, sino el análisis situacional.* El análisis que el historiador hace de la situación no es más que una conjetura histórica que, en este caso, constituye una metateoría acerca del razonamiento del emperador. Puesto que está a un nivel distinto del razonamiento del emperador, no lo evoca, sino que trata de reconstruirlo razonada e idealmente, omitiendo elementos esenciales y, quizá, ampliándolos. Por tanto, el metaproblema central del historiador es el siguiente: ¿Cuáles eran los elementos decisivos de la situación problemática del emperador? En la medida en que el historiador consiga resolver este metaproblema, *comprenderá* la situación histórica.

Por tanto, *qua* historiador, lo que tiene que hacer no es evocar experiencias pasadas, sino disponer argumentos objetivos en pro y en contra de su análisis conjetural de la situación.

Dicho método puede tener un éxito completo incluso en aquellos casos en que cualquier intento de evocación está llamado a fracasar, puesto que puede haber hechos que estén más allá de la capacidad de acción y, por tanto, de evocación del historiador. El acto a evocar puede ser de una crueldad insoportable, de supremo heroísmo o de cobardía despreciable. Tal vez se trate de un logro artístico, literario, científico o filosófico de tal elevación que exceda con mucho la capacidad del historiador. Está claro que si su competencia en el dominio que intenta analizar es insuficiente, sus análisis carecerán de interés. Ahora bien, no podemos esperar (como hace Colling-

---

[40] Cf. R. G. Collingwood, *The Idea of History,* 1946, pág. 283. (El subrayado es mío.) [Versión castellana, *Idea de la Historia,* México, Fondo de Cultura Económica, 2.ª ed., 1965, pág. 272.]

wood) que el historiador reúna las dotes de un César, un Cicerón, un Catullo y un Teodosio. Ningún historiador del arte podrá ser un Rembrandt, incluso pocos habrá que sean capaces de copiar una gran obra maestra.

Si bien, en los casos más interesantes, el historiador será incapaz de llevar a cabo la evocación, en otros podrá hacerlo, aunque sea algo completamente superfluo. Me refiero a esos innumerables casos triviales en los que es obvio, una vez que se ha analizado la situación, que la acción del agente era adecuada a la situación de un modo trivial y ordinario.

Por tanto, la misión del historiador es reconstruir la situación del problema tal como se presentaba al agente, de manera que sus actos se hagan adecuados a dicha situación. Este método se asemeja mucho al de Collingwood, si bien elimina de la teoría de la comprensión y del método histórico precisamente el elemento subjetivo o del segundo mundo que constituye el aspecto más destacado de la teoría (hermenéutica) para Collingwood y la mayoría de los teóricos de la comprensión.

Nuestra reconstrucción conjetural de la situación puede constituir un descubrimiento histórico real. Puede explicar un aspecto de la historia hasta el presente inexplicado y puede verse corroborada por nuevos elementos de juicio; por ejemplo, por el hecho de mejorar nuestra comprensión de un documento, llamando nuestra atención, pongo por caso, acerca de algunas alusiones que anteriormente habían pasado desapercibidas o habían quedado sin explicar [41].

---

[41] A partir de la teoría de las mareas de Galileo y sus realizaciones con Kepler, ya discutidas, hay otro ejemplo de interpretación que se podría mencionar aquí. En las páginas 103-5 de *Conjectures and Refutations* [trad. cit., págs. 20-2] he discutido la "*interpretario naturae*" de Bacon, señalando que significa "la lectura o exposición del libro de la naturaleza" y que el término "*interpretatio*" poseía un sentido legal distinto del moderno: en Bascon significa leer y "exponer" la ley (al profano) *tal como es exactamente*. (Esta interpretación mía surge plenamente del *De Argumentis*, lcc. cit., y arroja mucha luz sobre *todo* este pasaje del libro, y no solamente sobre la parte citada aisladamente en una nota anterior.) En el mismo lugar de mi *Conjectures*, también explico la idea baconiana relativa a la pureza y purificación del intelecto: significa purgar el intelecto de prejuicios, es decir, de teorías (de *anticipationes mentis*).

Ahora bien, sucede que Dilthey (*Schriften*, vol. V. pág. 318) interpreta erróneamente la "*interpretatio naturae*" de Bacon, que toma equivocadamente por una metáfora (ya que la toma en el sentido moderno de "interpretación" cuyo significado es casi el mismo que el de "*anticipatio mentis*" de Bacon). De manera similar, Ranke (*Sämtliche Werke*, vol. 49, pág. 175) malinterpreta la idea baconiana de pureza. Es evidente que si aceptamos mi interpretación conjetural, teniendo en cuenta el contexto, en el pasaje de Bacon discutido por Ranke, el autor (que escribe en latín) emplea "*caste*" en el sentido de "modestamente" (en el sentido intelectual, como muestra el contexto, de no incurrir en anticipaciones o declaraciones oraculares). Sin embargo, Ranke traduce equivocadamente "*caste*" por "honestamente" ("*keusch*"). Paralelamente, la expresión baconiana "*caste et perpetuo*" significa "modesta y constantemente" (podría traducirse libremente por "modesta y devotamente") y no "honesta y laboriosamente" ("*keusch und fleissig*"), como traduce equivocadamente Ranke. También hay en Ranke una atribución equivocada de esta mala traducción que me resulta totalmente inexplicable. Ascribe este pasaje al "Prólogo del *Organum*, con seguridad uno de los más bellos *Prefacios* que se hayan escrito nunca". ¿Cuáles son los hechos? El *Novum Organum* tiene un *Prefacio* en el que no aparece el pasaje que cita Ranke: aparece en el *Prefacio* de la *Instauratio Magna*, publicada junto con el *Organum*, aunque separada de su Prefacio por más de una docena de páginas (la *Distributio*

En resumen, he intentado mostrar que la idea del tercer mundo tiene interés para una teoría de la comprensión que trate de conjugar la comprensión intuitiva de la realidad con la objetividad de la crítica racional.

Operis, junto con una breve explicación en la que se advierte que falta la primera parte de la *instauratio*).

El pasaje puede traducirse del modo siguiente. (El texto está en la pág. 130, "*Nos vero...*" del vol. I de *The Works of Francis Bacon,* editadas por J. Spedding, R. L. Ellis y D. D. Heath, 1889.) "Pero yo, permaneciendo modesta y constantemente entre las cosas [mismas], nunca aparto el intelecto de ellas más allá de lo preciso para permitir enfocar sus imágenes y rayos, como es posible hacerlo con el sentido [de la vista]". (Después de poner punto y coma, Bacon termina diciendo: "de modo que no tengan mucho que hacer las facultades de invención y superación".)

He aquí la traducción y comentarios de Ranke: " 'Lasst uns', sagt Bacon in der Vorrede zu dem Organon —gewiss einem der schönsten Prooemien, die je geschrieben worden sind— 'lasst uns keusch und fleissig unter den Dingen verweilen und unsere Fassungskraft nur eben so weit über sie erheben, um ihre Bilder und Strahlen in uns aufhehmen zu können'.

"Er sagte dies von der Betrachtung der *Natur.* Die Erforschung der *Geschichte* hat es freilich noch schwerer". (*) (Etc.: Ranke se ocupa de las dificultades especiales de la historiografía —de la interpretación de la *historia* frente a la interpretación de la *naturaleza.*)

Como puede verse por la mala traducción que hace Ranke del sencillo texto latino de Bacon, la interpretación (*hermenéutica*) de textos que, después de todo, forma parte de la historiografía, es casi tan arriesgada como la interpretación de la naturaleza Constituye un terreno en el que hemos de proceder mediante conjeturas y refutaciones: es decir, hemos de intentar refutar nuestras conjeturas hasta que encajen plenamente en el contexto de la situación problemática, se desprendan de sus aspectos arbitrarios y consigan algo así como el máximo poder explicativo de lo que el autor quería decir.

Para otros ejemplos del método conjetural de interpretación, véanse especialmente las notas al primer volumen de mi *Open Society* y los *Apéndices* 6 a 9 de mi *Conjectures and Refutations,* tercera edición, 1969 y cuarta edición, 1972.

(*) " 'Dejadnos", dice Bacon en el Prólogo del Organon —sin duda uno de los Prefacios más bellos que se hayan escrito nunca—, 'dejadnos permanecer honesta y laboriosamente en las cosas y elevar sobre ellas nuestras facultades intelectuales tan sólo para poder asumir en nosotros sus imágenes y destellos'.

"Decía esto de la contemplación de la *Naturaleza.* La exploración de la *Historia* es un problema aún más difícil". [*N. T.*].

5. EL OBJETO DE LA CIENCIA*

Puede parecer un poco ingenuo hablar del "objeto de la ciencia", ya que, como es evidente, distintos científicos tendrán diferentes metas y la propia ciencia (sea lo que sea lo que esto quiera decir) carece de metas. Aunque admito todo esto, no obstante, me parece que cuando hablamos de ciencia nos da la sensación, más o menos clara, de que hay algo característico de la actividad científica. Por tanto, si la actividad científica se asemeja tanto a la actividad racional y ésta ha de tener alguna meta, entonces no será completamente fútil intentar describir el objeto de la ciencia.

Sugiero que el objeto de la ciencia consiste en dar con *explicaciones satisfactorias* de todo aquello que nos parece precisar una explicación. Por *explicación* (o explicación causal) se entiende un conjunto de enunciados mediante los cuales se describe el estado de la cuestión a explicar (el *explicandum*) sirviéndose para ello de otros, los enunciados explicativos, que constituyen la "explicación" en sentido estricto (el *explicans* del *explicandum*).

Por regla general, damos por descontado que sabemos más o menos que el *explicandum* es verdadero o suponemos que lo es, ya que carece de interés pedir una explicación de algo que puede resultar totalmente imaginario. (Los platillos volantes son un buen ejemplo: la explicación requerida no versa acerca de ellos, sino acerca de los *informes* que hablan de ellos; con todo, si existiesen los platillos volantes, entonces no se necesitaría explicar además los *informes*.) Por otro lado, por regla general, no conoceremos el *explicans* que constituye el objeto de nuestra investigación: habrá de ser descubierto. Por tanto, la explicación científica, en la medida en que constituya un descubrimiento, será la *explicación de lo conocido mediante lo desconocido* [1].

Para que el *explicans* sea satisfactorio (la satisfactoriedad es cuestión de

---

* Este artículo es una versión revisada del publicado en *Ratio*, vol. I, núm. 1, diciembre 1957, págs. 24-35. Un breve comentario sobre la corrección de los resultados de Galileo y Kepler mediante la teoría newtoniana apareció publicado por primera vez en mi contribución al libro de Simon Moser (ed.), *Gesetz und Wirklichkeit*, 1949 (véanse concretamente las págs. 57 y sigs.), reimpresa en Hans Albert, *Theorie und Realität*, 1964 (especialmente, pág. 100). En el Apéndice de este mismo volumen se puede encontrar una traducción al inglés de dicho escrito.

[1] Véase el último párrafo del texto, antes de la cita final, de mi "Nota sobre Berkeley como Precursor de Mach", *Brit. Journ. Philos. Sc., 4*, 1953, pág. 35. (Incluido ahora en mi libro *Conjectures and Refutations*, pág. 174 [traducción castellana de Néstor Míguez, *El Desarrollo del Conocimiento Científico. Conjeturas y Refutaciones*, Buenos Aires. Paidos, 1967, pág. 203]).

grado) tienen que reunir ciertos requisitos. En primer lugar, ha de implicar lógicamente el *explicandum* En segundo lugar, el *explicans* ha de ser verdadero, aunque en general no lo sepamos; en cualquier caso no hemos de saber que es falso hasta después del examen más crítico. Si no sabemos que es verdadero (que será lo más corriente), habremos de disponer de elementos de juicio *independientes* en favor suyo. En otras palabras, debe ser contrastable *independientemente;* además, lo consideraremos como más satisfactorio cuanto más rigurosas sean las contrastaciones independientes de las que haya salido indemne.

Así pues, tenemos que aclarar en qué sentido empleo la expresión "independiente" y sus opuestas, "*ad hoc*" y (en casos extremos) "circular".

Sea *a* un *explicandum* que sabemos que es verdadero. Puesto que *a* se sigue del propio *a* de un modo trivial, siempre podemos ofrecer *a* como explicación de sí mismo. Ahora bien, se trata de algo totalmente insatisfactorio, aunque en este caso sepamos que el *explicans* es verdadero, y que de él se sigue el *explicandum. Por tanto hemos de rechazar este tipo de explicaciones por su carácter circular.*

Sin embargo, el tipo de circularidad que me preocupa más es una cuestión de grado. Tomemos el siguiente diálogo: "¿Por qué está hoy el mar tan encrespado?" —"Porque Neptuno está furioso"— "¿En qué se basa usted para decir que Neptuno está furioso?" —"Caramba, ¿no *ve* usted *qué* encrespado está el mar? ¿Acaso no se encuentra así cuando Neptuno está furioso?" Consideramos insatisfactoria esta explicación porque (como en el caso de las explicaciones plenamente circulares) el único elemento de juicio en pro del *explicans* es el propio *explicandum* [2]. La sensación que produce este tipo de explicación casi circular o *ad hoc* es muy *insatisfactoria,* por lo que creo que el requisito correspondiente de que hemos de evitarlas es una de las fuerzas motoras más importantes del desarrollo científico: la insatisfacción es uno de los primeros frutos del enfoque crítico o racional.

Para que el *explicans* no sea *ad hoc* habrá de tener un contenido muy rico: habrá de tener una gran variedad de consecuencias contrastables, entre las que habrán de encontrarse, especialmente, algunas distintas del *explicandum.* Cuando hablo de contrastaciones *independientes* o de elementos de juicio *independientes* me refiero precisamente a esas consecuencias contrastables distintas.

Aunque estas consideraciones tal vez ayuden a elucidar en alguna medida la idea intuitiva de *explicans* contrastable independientemente, siguen siendo insuficientes para caracterizar una explicación satisfactoria contrastable independientemente. La razón estriba en que si *a* es nuestro *explicandum* —sea *a* otra vez, "Hoy el mar está encrespado"— siempre podemos suministrar un *explicans* totalmente insatisfactorio y *ad hoc,* aún cuando tenga consecuencias contrastables independientemente. Podemos elegir como

---

[2] Este tipo de razonamiento aparece en Tales (Diels-Kranz [10], vol. I, pág. 456, línea 35), en Anaximandro (D.-K. All, B28), en Anaxímenes (D.-K. A17, B1) y en Alcmeón (D.-K. A5).

queramos esas consecuencias. Tomemos, por ejemplo, "Estas ciruelas son jugosas" y "Todos los cuervos son negros". Sea *b* su conjunción. A continuación podemos tomar como *explicans* sencillamente la conjunción de *a* y *b* que satisfará todos los requisitos exigidos hasta ahora.

Solamente si exigimos que las explicaciones utilicen enunciados universales o leyes de la naturaleza (acompañadas de condiciones iniciales), podremos avanzar algo hacia la idea de explicación independiente y no *ad hoc*. Es así, porque las leyes universales de la naturaleza *han* de ser enunciados que posean un contenido rico, de manera que *puedan ser contrastadas independientemente* en todo tiempo y lugar. Por tanto, si las utilizamos como explicaciones, no *podrán ser ad hoc*, puesto que nos permiten interpretar el *explicandum* como un caso particular de un efecto reproductible. Ahora bien, todo esto sólo vale en caso de que nos limitemos a leyes universales que sean contrastables, es decir, falsables.

La pregunta "¿Qué tipo de explicación puede considerarse satisfactorio?", nos lleva a responder: una explicación en términos de leyes universales contrastables y falsables junto con condiciones iniciales. Además, una explicación tal será tanto más satisfactoria, cuanto más altamente contrastables sean estas leyes y cuanto mejor contrastadas hayan sido. (Esto también sirve para las condiciones iniciales.)

De este modo, la conjetura según la cual el objeto de la ciencia es dar con explicaciones satisfactorias nos conduce a la idea de mejorar el grado de satisfactoriedad de las explicaciones mejorando su grado de contrastabilidad; es decir, yendo hacia teorías que se puedan contrastar mejor o, lo que es lo mismo, yendo hacia teorías que tengan un contenido cada vez más rico, un grado de universalidad cada vez más elevado y un grado de precisión cada vez mayor [3]. No cabe duda de que esto está plenamente de acuerdo con la práctica real de las ciencias teóricas.

También podemos llegar por otro camino al mismo resultado. Si el objeto de la ciencia es explicar, también tendrá como misión explicar lo que hasta ahora se ha tomado como *explicans;* por ejemplo, una ley de la naturaleza. Por tanto, la tarea científica se renueva constantemente. Podemos avanzar eternamente, procediendo a dar explicaciones con un grado de universalidad más y más elevado, a menos que lleguemos a *explicaciones últimas;* es decir, a una explicación que ni necesita ni es susceptible de una explicación ulterior.

Ahora bien, ¿acaso hay explicaciones últimas? La doctrina que he denominado "esencialismo" viene a decir que la ciencia ha de buscar explicaciones últimas en términos de esencias [4]: si podemos explicar el comportamiento

---

[3] Para la teoría de la *contrastabilidad, contenido, simplicidad* y grados de *universalidad y precisión,* véanse las secciones 31-46 de mi libro *Logic of Scientific Discovery,* 1956 (primera edición alemana, 1934; cuarta edición alemana, 1971 [traducción castellana de Víctor Sánchez de Zavala, *La Lógica de la Investigación Científica.* Madrid, Tecnos, 1962].) En el que se explica la íntima relación existente entre estas ideas.

[4] He discutido (y criticado) el esencialismo de un modo más extenso en mi artículo "Tres Concepciones Sobre el Conocimiento Humano", donde también hago

de algo en términos de su esencia —de sus propiedades esenciales—, entonces, no pueden plantearse más preguntas, ni falta que hace (excepto, quizá, la cuestión teológica del creador de las esencias). Así Descartes creía haber explicado la física en términos de la *esencia del cuerpo físico* que, según creía, era la extensión. Algunos newtonianos, seguidores de Roger Cotes, pensaban que la *esencia de la materia* era la inercia y su capacidad de atraer otras materias, con lo cual la teoría newtoniana se podía derivar y, en consecuencia, explicar últimamente, mediante esas propiedades esenciales de toda materia. El propio Newton era de otra opinión. Tenía en la mente una hipótesis relativa a la explicación causal última o esencialista de la gravedad misma cuando escribió en el *Scholium Generale*, al final de los *Principia*: "Hasta ahora he explicado los fenómenos... mediante la fuerza de la gravedad, pero hasta el momento no he averiguado la *causa de la gravedad misma*... y no invento hipótesis arbitrariamente (*ad hoc*)" [5].

No creo en la doctrina esencialista de las explicaciones últimas. Por regla general, los críticos de esta doctrina en tiempos pasados han sido instrumentalistas: interpretaban las teorías científicas *únicamente como* instrumentos de predicción carentes de poder explicativo. Tampoco estoy de acuerdo con ellos. Sin embargo, hay una tercera posibilidad, un tercer "punto de vista", como lo he dado en llamar. Se puede caracterizar perfectamente como un "esencialismo modificado", haciendo hincapié en la palabra "modificado" [6].

El "tercer punto de vista" que yo mantengo modifica de manera radical el esencialismo. Ante todo, rechaza la idea de explicación última: sostengo que toda explicación puede ser explicada a su vez mediante una teoría o conjetura de un grado superior de universalidad. No hay explicaciones que no precisen ulteriores explicaciones, porque ninguna puede constituir una descripción auto-explicativa de una esencia (como ocurre con la definición esencialista de cuerpo que apunta Descartes). En segundo lugar, rechazo las *preguntas* del tipo ¿Qué es?; es decir, preguntas acerca de lo que es una cosa, acerca de cuál es su esencia o su verdadera naturaleza. La causa de ello estriba en que hemos de rechazar la opinión, típica del esencialismo, de que

---

alusión a mis discusiones anteriores (en la última nota de la sección II); véase (*Contemporay British Philosophy*, III, editado por H. D. Lewis, 1956, nota 2 de la pág. 365. (Este artículo constituye ahora el capítulo 3 de mi libro *Conjectures and Refutations* [trad. cit., págs. 116-41], 3.ª edición, 1969.)

[5] Véase también la carta de Newton a Bentley del 17 de enero y especialmente la del 25 de febrero de 1693 ("1692-3"). He sacado citas de esta carta en la sección III de mi artículo "Tres Concepciones Sobre el Conocimiento Humano" (*Conjectures and Refutations*, págs. 106 y sig. [trad. cit., pág. 127]), donde discuto el problema un poco más detenidamente.

[6] La expresión "esencialismo modificado" la ha empleado, para describir mi "tercera posición", el autor de una recensión de mi artículo "Tres Concepciones Sobre el Conocimiento Humano", aparecida en *The Times Literary Supplement*, 55, 1956, pág. 527. Para evitar malentendidos, he de decir aquí que el que yo acepte esta expresión no ha de interpretarse como una concesión a la doctrina de la "realidad última" ni mucho menos como una concesión a la de las definiciones esenciales. Estoy plenamente de acuerdo con la crítica a esta doctrina expuesta en mi *Open Society*, vol. II, cap. III, sección II (especialmente, la nota 42) [traducción castellana de Eduardo Loedel, *La Sociedad Abierta y sus Enemigos*, Buenos Aires, Paidos, 1957], así como en otros lugares.

cada cosa tomada aisladamente posee una esencia, una naturaleza o principio inherente (como el espíritu del vino en el vino), que provoca necesariamente que sea lo que es y actúe como lo hace. Esta posición animista no explica nada, si bien ha llevado a los esencialistas (como Newton) a rehuir las propiedades relacionales, como la gravedad, y a creer, basándose en fundamentos que creían válidos *a priori*, que una explicación satisfactoria había de formularse en términos de propiedades inherentes (frente a las relacionales). La tercera y última modificación del esencialismo es la siguiente; hemos de abandonar la opinión, íntimamente relacionada con el animismo (y típica de Aristóteles frente a Platón) de que para explicar el comportamiento de una cosa hemos de recurrir a las propiedades esenciales inherentes *a cada cosa individual o singular*. La razón es que esta opinión no logra en absoluto arrojar ninguna luz sobre el problema de por qué distintas cosas individuales se comportan de manera semejante. Si se dice que "porque sus esencias son semejantes", surge una nueva pregunta *¿por qué no podría haber tantas esencias diferentes como cosas distintas hay?*.

Platón intentó resolver precisamente este problema diciendo que la semejanza de las cosas individuales se deriva de (y, por tanto, copia) la misma "Forma" original que es, en consecuencia, "previa", "exterior" y "superior" a las diversas cosas individuales. Además, no disponemos de una teoría mejor sobre la semejanza. Incluso hoy, para explicar la semejanza entre dos hombres, pájaros, peces, camas, coches, lenguajes o procedimientos legales recurrimos a su origen común; es decir, explicamos la semejanza de un modo fundamentalmente genético y si, a partir de ahí, construimos un sistema metafísico, lo más fácil es que se convierta en una filosofía historicista. Aristóteles no aceptó la solución platónica, pero, puesto que la versión aristotélica del esencialismo no contiene el menor rastro de solución, da la impresión de que nunca llegó a comprender plenamente el problema [7].

La elección de explicaciones en términos de leyes naturales universales constituye una solución precisamente a este último problema (platónico), ya que consideramos que todas estas cosas individuales y todos estos hechos singulares están sujetos a esas leyes. Por tanto, las leyes (que a su vez *precisan* una explicación ulterior) explican las regularidades o semejanzas de las cosas individuales y de los hechos o eventos singulares. Además, dichas leyes no son inherentes a esas cosas singulares (ni las ideas platónicas están fuera del mundo). Las leyes de la naturaleza las pensamos, más bien, como descripciones (conjeturales) de las propiedades estructurales de la naturaleza —o de nuestro propio mundo.

Aquí es donde hay que buscar, pues, la semejanza entre mi punto de vista (la "tercera concepción") y el esencialismo. Aunque no creo que mediante leyes universales podamos llegar a describir la esencia *última* del mun-

---

[7] Por lo que atañe a la teoría platónica de las Formas o Ideas, "una de sus funciones más importantes" es "explicar la semejanza de las cosas sensibles..." Cf. mi *Open Society*, cap. 3, sección V; véanse también las notas 19 y 20, así como el texto. Menciono allí el fracaso de la teoría aristotélica a la hora de desempeñar esta función (en la tercera edición de 1957) al final de la nota 57 del capítulo II.

do, no me cabe la menor duda de que podemos esforzarnos por penetrar cada vez con mayor profundidad en la estructura de nuestro mundo o, podríamos decir, en las propiedades del mundo cada vez más esenciales o de mayor grado de profundidad.

Cada vez que procedemos a explicar alguna ley o teoría conjetural, sirviéndonos de una nueva teoría de un grado de universalidad superior, descubrimos más cosas acerca del mundo, al intentar profundizar más en sus secretos. Además, cada vez que logramos falsar una teoría de este tipo, realizamos un nuevo descubrimiento importante. Dichas falsaciones son de la mayor trascendencia porque nos enseñan algo inesperado y nos reafirman en la idea de que, aunque las teorías las hagamos nosotros, aunque sean un invento hecho por nosotros, no obstante, constituyen genuinas afirmaciones acerca del mundo, ya que pueden *chocar* con algo que no hemos hecho nosotros.

Creo que nuestro "esencialismo modificado" resulta útil a la hora de plantear el problema de la forma lógica de las leyes naturales. Pienso que nuestras leyes o teorías han de ser *universales,* es decir, deben afirmar algo acerca del mundo —acerca de todas las regiones espacio— temporales del mundo. Afirmo además, que nuestras teorías hablan acerca de las propiedades estructurales o relacionales del mundo y que las propiedades descritas por una teoría explicativa tiene que ser, en algún sentido, más profundas que aquellas que hay que explicar. Creo que la expresión "más profundo" desafía cualquier intento de análisis lógico exahustivo, si bien, a pesar de todo, guía nuestra intuición. (Es algo que ocurre en el campo de las matemáticas: todos sus teoremas son lógicamente equivalentes respecto a los axiomas, aún cuando haya una gran diferencia en "profundidad" difícilmente susceptible de análisis lógico.) La "profundidad" de una teoría científica parece estar íntimamente relacionada con su simplicidad y, consiguientemente, con la riqueza de su contenido. Se trata de algo distinto de la profundidad de un teorema matemático cuyo contenido puede decirse que es nulo. Al parecer, se requieren dos ingredientes: un contenido rico y una cierta coherencia o un cierto carácter compacto (u "organicidad") del estado de la cuestión descrito. Es éste último ingrediente, que resulta tan difícil de analizar a pesar de su gran claridad intuitiva, el que los esencialistas intentaban describir al hablar de esencias en contraposición a una mera acumulación de propiedades accidentales. No creo que aquí podamos hacer muho más que remitirnos a una idea intuitiva ni que se precise más que eso, ya que en el caso de una teoría particular que se presente ante nosotros, lo que decide acerca de su interés es la riqueza de su contenido y, por tanto, su grado de contrastabilidad, del mismo modo que lo que decide acerca de su destino es el resultado de las contrastaciones efectivas. Desde un punto de vista metodológico, hemos de considerar su profundidad, su coherencia e incluso su belleza como una simple guía o estímulo para nuestra intuición e imaginación.

No obstante, parece haber algo así como una condición *suficiente* para la profundidad o para los grados de profundidad susceptible de análisis ló-

gico. Intentaré explicarlo mediante un ejemplo sacado de la historia de la ciencia.

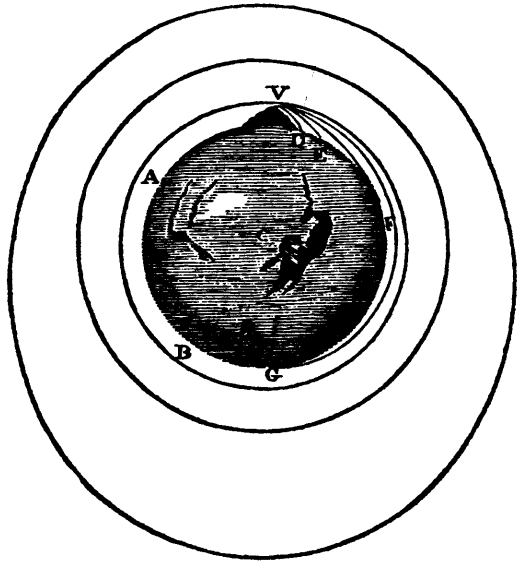
Es bien sabido que la dinámica newtoniana llevó a cabo una reunificación de la física terrestre de Galileo y de la celeste de Kepler. A veces se dice que la dinámica de Newton pude inducirse a partir de las leyes de Galileo y Kepler, e incluso se ha llegado a afirmar que se puede deducir de ellas en sentido estricto [*]. Pero no es ese el caso; desde un punto de vista lógico, la teoría newtoniana contradice, estrictamente hablando, tanto a la de Galileo como a la de Kepler (si bien, como es obvio, tan pronto como disponemos de la teoría de Newton, pueden ser obtenidas ambas a título de aproximaciones). Por esta razón es imposible derivar la teoría newtoniana sea de la de Galileo, sea de la de Kepler, sea de ambas a la vez, sirviéndonos para ello de la inducción o de la deducción. La razón estriba en que ni una inferencia inductiva ni una deductiva puede pasar de premisas consistentes a una conclusión que contradiga formalmente las premisas de que hemos partido.

Creo que esto constituye un argumento poderoso en contra de la inducción.

Indicaré ahora brevemente las contradicciones existentes entre la teoría newtoniana y las de sus predecesores. Galileo afirma que una piedra lanzada al aire o un proyectil se mueven según una parábola, excepto en el caso de la caída libre vertical, en el cual se mueven en línea recta con aceleración constante. (En esta discusión prescindimos de la resistencia del aire.) Desde el punto de vista de la teoría newtoniana, ambas afirmaciones son falsas por dos razones. La primera es falsa porque la trayectoria de un proyectil de largo alcance, como puede ser un cohete intercontinental (lanzado horizontalmente o hacia arriba) no será parabólica, sino elíptica. Sólo se aproxima a una parábola si la distancia total de vuelo del proyectil es despreciable respecto al radio terrestre. El propio Newton señaló este punto en sus *Principia*, así como en su versión popular, *The System of the World*, donde lo ilustra mediante la figura reproducida aquí.

---

[*] Lo que se puede deducir de las leyes de Kepler (Véase Max Born, *Natural Philosophy of Cause and Chance*, 1949, págs. 129-33) es que para todos los planetas, la aceleración hacia el Sol es en todo momento igual a k/r^2, siendo r la distancia que separa al Sol del planeta en ese momento y, k una constante para todos los planetas. Sin embargo, este último resultado contradice la teoría newtoniana (a menos que supongamos que todas las masas de los planetas son iguales o que, en caso de ser desiguales, lo sean en una proporción infinitamente pequeña comparada con la masa del Sol). Este hecho se deriva de lo que aquí decimos (en la siguiente nota, 9) acerca de la tercera ley de Kepler. Pero, además, habrá que recordar que ni la teoría de Galileo ni la de Kepler contienen el concepto newtoniano de *fuerza* que se incluye tradicionalmente en estas deducciones sin más problemas, como si dicho concepto ("oculto") se pudiese leer en los hechos en lugar de ser el resultado de una nueva interpretación de ellos (es decir, de los "fenómenos" descritos por las leyes de Kepler y Galileo) a la luz de una teoría totalmente nueva. Sólo después de haber introducido el concepto de fuerza (e incluso la proporcionalidad entre las masas inercial y gravitatoria) es posible relacionar la fórmula anterior de la aceleración con la ley newtoniana del inverso del cuadrado de la distancia (suponiendo que las masas de los planetas sean despreciables).



La figura de Newton ilustra su enunciado según el cual, si la velocidad del proyectil crece y con ella las distancias recorridas durante el vuelo, "acabará superando los límites de la tierra..., y pasará al espacio sin tocarla" *

Por tanto, un proyectil terrestre se mueve siguiendo una elipse más bien que una parábola. Evidentemente, para lanzamientos suficientemente cortos, la parábola constituye una aproximación excelente. Con todo, en sentido estricto, de la teoría de Newton no puede deducirse una trayectoria parabólica, a menos que se añada una ulterior condición inicial, de hecho *falsa,* (y, dicho sea de paso, impracticable en la teoría newtoniana porque lleva a consecuencias absurdas) según la cual el radio de la tierra es infinito. Si no aceptamos dicha suposición que *sabemos que es falsa,* obtendremos siempre una elipse, en contra de la ley de Galileo según la cual obtenemos siempre una parábola.

Se plantea una situación lógica análoga en relación con la segunda parte de la ley de Galileo, según la cual se da una aceleración *constante.* Desde el punto de vista de la teoría newtoniana, la aceleración de los cuerpos en caída libre nunca es constante: aumenta paulatinamente durante la caída, puesto que el cuerpo se aproxima cada vez más al centro de atracción. El efecto es de magnitud considerable cuando el cuerpo cae de una altura muy grande, si bien es despreciable cuando la altura es muy pequeña en relación al radio terrestre. En este caso podemos obtener la teoría de Galileo a partir de la de

---

* Véase el *Scholium* que aparece al final de la sección II del Libro I de los *Pricipia* de Newton; pág. 55 de la edición de 1934 (traducción de Motte revisada por Cajori). Tanto la figura de *The System of the World* como la cita se encontrarán en la pág. 591 de esta edición.

Newton si introducimos, una vez más, la suposición *falsa* de que el radio de la tierra es infinito (o la altura de la caída, cero).

Las contradicciones que he señalado no son en absoluto despreciables en caso de proyectiles de largo alcance. A éstos hemos de aplicarles la teoría newtoniana (evidentemente, teniendo en cuenta las correcciones necesarias debidas a la resistencia del aire) y no la de Galileo que lleva sencillamente a resultados falsos, como puede verse claramente con ayuda de la teoría de Newton.

La situación es similar por lo que respecta a las leyes de Kepler. Es obvio que para la teoría newtoniana las leyes de Kepler no son más que aproximadamente válidas —es decir, estrictamente inválidas— si tenemos en cuenta las atracciones entre los planetas [10]. Ahora bien, entre ambas teorías hay contradicciones más fundamentales que las derivadas de este hecho obvio. En efecto, aunque como concesión a nuestros oponentes despreciemos las atracciones mutuas entre los planetas, la tercera ley de Kepler, considerada desde el punto de vista de la dinámica newtoniana, no puede ser más que una aproximación aplicable a un caso muy especial: a los planetas de igual masa o, de ser sus masas desiguales, despreciables en relación a la masa del sol. Puesto que esto no es válido, ni siquiera aproximadamente si tomamos dos planetas, uno muy ligero y otro muy pesado, es evidente que la tercera ley de Kepler contradice la teoría newtoniana exactamente del mismo modo que ocurre con la de Galileo.

Esto puede verse fácilmente del siguiente modo. En el caso de un sistema de dos cuerpos —un sistema binario de estrellas— la teoría de Newton presenta una ley llamada a veces por los astrónomos "ley de Kepler" por su estrecha relación con la tercera ley de Kepler. Esta llamada "ley de Kepler" dice que si m_0 es la masa de uno de los dos cuerpos —el sol, digamos— y m_1 la del otro cuerpo —un planeta, por ejemplo— entonces tomando unidades de medición adecuada podemos derivar de la teoría de Newton:

$$(1) \qquad a^3/T^2 = m_0 + m_1,$$

siendo a la distancia media entre ambos cuerpos y T, el tiempo de una revolución completa. Ahora bien, la ley tercera del propio Kepler afirma que:

$$(2) \qquad a^3/T^2 = \textit{constante;}$$

es decir, la misma constante para *todos* los planetas del sistema solar. Está claro que (2) sólo se obtiene a partir de (1) si suponemos que $m_0 + m_1 =$ cons-

[10] Véase, por ejemplo, P. Duhem, *The Aim and Structure of Physical Theory*, 1905; versión inglesa de P. P. Wiener, 1945, Parte II, capítulo VI, sección 4. Duhem formula más explícitamente lo que estaba implícito en el propio enunciado de Newton (*Pricipia*, Libro I, proposición LXV, teorema XXV), ya que Newton deja bien claro que en los casos en que entran en interacción más de dos cuerpos, las leyes de Kepler serán, a lo sumo, tan sólo aproximadamente válidas y esto, únicamente en casos muy especiales, dos de los cuales analiza con algún detalle. Dicho sea de paso, la fórmula (1) que aparece más adelante en el texto se sigue inmediatamente de la proposición LIX del libro I con ayuda de la proposición XV del mismo Libro (Véase también la proposición XV del Libro III.)

tante; y puesto que m_0 = constante para nuestro sistema solar, si identificamos m_0 con la masa del sol, obtendremos (2) a partir de (1) siempre y cuando supongamos que m_1 es igual para todos los planetas; o bien, puesto que de hecho esto es *falso* (como es el caso, ya que Júpiter es varios miles de veces mayor que el menor de los planetas), que las masas de los planetas sean *todas cero por respecto a la del sol,* de modo que podamos escribir $m_1 = 0$, *para todos los planetas.* Es una buena aproximación desde el punto de vista de la teoría newtoniana; pero, al mismo tiempo, escribir $m_1 = 0$, no sólo es falso estrictamente hablando, sino además, impracticable desde el punto de vista de la teoría newtoniana. (Un cuerpo con masa cero no obedecería las leyes del movimiento de Newton.) Por tanto, aunque desestimemos todo lo referente a las atracciones mutuas entre los planetas, la tercera ley de Kepler (2) contradice la teoría newtoniana de la que se deriva (1).

Es importante constatar que, partiendo de las teorías de Galileo o de Kepler, no recibimos la menor sugerencia acerca de cómo ajustar dichas teorías —qué falsas premisas adoptar o qué condiciones estipular— en orden a pasar a otra teoría con validez más general, como ocurre con la de Newton. *Sólo después de entrar en posesión de la teoría de Newton podemos descubrir si, y en qué sentido, las viejas teorías pueden considerarse como aproximaciones suyas.* Podemos expresar esto brevemente diciendo que, aunque las teorías de Galileo y Kepler sean, desde el punto de vista de la newtoniana, excelentes aproximaciones a determinados resultados especiales de Newton, no puede decirse que su teoría sea una aproximación, desde el punto de vista de las otras dos, a sus resultados. Esto muestra que la lógica, sea deductiva o inductiva, no puede dar el paso que media entre estas teorías y la dinámica newtoniana [11]. Sólo el ingenio puede dar ese paso. Una vez dado, puede decirse que los resultados de Kepler y Galileo corroboran la nueva teoría.

No obstante, lo que ahora me preocupa no es tanto el problema de la inducción, cuanto *el de la profundidad.* Si tenemos en cuenta este problema, podemos sacar algo en limpio del ejemplo. La teoría de Newton unifica las de Galileo y Kepler. Ahora bien, lejos de ser una simple conjunción de ambas teorías —que para Newton desempeñan el papel de *explicanda— las corrige, al tiempo que las explica.* La tarea explicativa original era la deducción de los resultados primitivos. Sin embargo, esta tarea se cumple, no deduciendo estos resultados primitivos, sino deduciendo en su lugar algo mejor: resultados nuevos que en condiciones especiales están muy próximos numéricamente a los viejos, a la vez que los corrigen. Así pues, puede decirse que el éxito empírico de una vieja teoría corrobora la nueva; además, las correcciones se pueden contrastar a su vez —tal vez refutándolas, tal vez corroborándolas—. La situación lógica que he bosquejado pone plenamente de manifiesto el hecho de que la nueva teoría no puede ser *ad hoc* o circular. Lejos

[11] Los conceptos de fuera (cf. la pág. 186, nota 8) y acción a distancia introducen aún más dificultades.

de limitarse a repetir el *explicandum*, la nueva teoría lo contradice y corrige. De este modo, los propios elementos de juicio en favor del *explicandum* se tornan en elementos de juicio independientes en favor de la nueva teoría. (Dicho sea de pasada, este análisis nos permite *explicar el valor de las teorías métricas* y de la medición, con lo que nos evita incurrir en el error de aceptar la medición y la precisión como valores últimos e irreductibles.)

Sostengo que cuando encontremos en las ciencias empíricas que una nueva teoría con un nivel superior de universalidad explica con éxito otra anterior, *corrigiéndola*, tenemos ahí una señal inequívoca de que la nueva teoría ha penetrado más profundamente que las anteriores. Siguiendo a Bohr, podemos denominar "*principio de correspondencia*" al requisito de que una nueva teoría contenga aproximadamente a la vieja para valores apropiados de los parámetros de la nueva. Satisfacer este requisito es una condición suficiente de profundidad, como ya he dicho. Que no constituye una condición necesaria se desprende del hecho de que la teoría ondulatoria electromagnética de Maxwell no corrigió, en este sentido, a la teoría ondulatoria de la luz de Fresnel. Sin duda significaba un aumento de profundidad, pero en otro sentido: "El viejo problema relativo a la dirección de las vibraciones de la luz polarizada perdió interés. Las dificultades relativas a las condiciones límite de la separación entre dos medios se vieron resueltas por los fundamentos mismos de la teoría. Ya no se necesitaban hipótesis *ad hoc* para eliminar las ondas lumínicas longitudinales. La presión de la luz, tan importante en la teoría de la radiación y que sólo pudo ser determinada experimentalmente más tarde, podía derivarse como una de las consecuencias de la teoría" [12]. Este brillante pasaje, en el que Einstein bosqueja algunos de los logros fundamentales de la teoría de Maxwell frente a la de Fresnel, puede tomarse como índice de la existencia de otras condiciones suficientes de profundidad que no estaban recogidas en mi análisis.

Si no somos realistas, difícilmente podremos comprender cuál es la tarea de la ciencia que, como he apuntado, consiste en dar con explicaciones satisfactorias. Una explicación satisfactoria es aquella que no es *ad hoc;* idea ésta —*la de elemento de juicio independiente*— que difícilmente se puede comprender sin recurrir a la de descubrimiento, a la de progreso hacia estratos más profundos de explicación: sin la idea de que hay algo que descubrir y discutir críticamente.

Sin embargo, me parece que en metodología no tenemos que presuponer un realismo metafísico ni creo que podamos sacar de él mucho en limpio, excepto a nivel intuitivo. La razón es que, una vez que se ha dicho que el objeto de la ciencia es explicar y que la explicación más satisfactoria será aque-

[12] A. Einstein, *Physikalische Zeitschrift, 10,* 1909, págs. 817 y sig. El abandono de la teoría del éter material (implícito en el fracaso de Maxwell a la hora de construir un modelo material satisfactorio) puede tomarse como un aumento de profundidad, en el sentido analizado arriba con ocasión de la comparación entre la teoría de Maxwell y la de Fresnel. Creo que esto aparece implícito en la cita del artículo de Einstein. Así, la teoría de Maxwell, tal como la formula Einstein, quizá no constituya realmente un ejemplo de *otro* sentido de "profundidad", si bien pienso que lo constituye si tomamos la versión original de Maxwell.

lla contrastable y contrastada de hecho de un modo más riguroso, sabemos todo lo que necesitamos saber como metodólogos. No podemos afirmar que esta meta sea realizable con o sin la ayuda del realismo metafísico que tan sólo nos puede proporcionar cierto estímulo intuitivo, ciertas esperanzas, pero no seguridad de ningún tipo. Aunque se pueda decir que el tratamiento racional de la metodología depende de lo que suponemos o conjeturamos que es el objeto de la ciencia, con todo, es evidente que no depende de la suposición metafísica y probablemente falsa de que la teoría estructural del mundo (si la hubiese) es susceptible de ser descubierta por los hombres o es expresable en lenguaje humano.

Si la imagen del mundo trazada por la ciencia moderna se acerca de algún modo a la verdad —en otras palabras, si poseemos algo así como un "conocimiento"— entonces las condiciones vigentes en casi todas partes del universo hacen casi imposible el descubrimiento de leyes estructurales del tipo buscado y, por consiguiente, el logro del "conocimiento científico". La razón de ello es que casi todas las regiones del universo están llenas de radiaciones caóticas y las regiones que quedan están casi todas ocupadas por materia en un estado de caos similar. A pesar de ello, la ciencia ha conseguido avanzar milagrosamente hacia aquello que, a mi modo de ver, constituye su meta. Creo que este hecho extraño no se puede explicar sin probar demasiado. Pero puede animarnos a perseguir esa meta, aún cuando no obtengamos un estímulo adicional (ni del realismo metafísico ni de ninguna otra parte) para creer que la hayamos alcanzado de hecho.

BIBLIOGRAFIA SELECTA

Señalo con asterisco los trabajos que aparecen en este volumen.

POPPER, KARL, R.: *Logik der Forschung*, 1934 (1935); ediciones aumentadas, 1966, 1969. [Traducción castellana de Víctor Sánchez de Zavala, *La Lógica de la Investigación Científica*. Madrid, Tecnos, 1962.]

— *The Poverty of Historicism* (1944-5), 1957, 1960. [Traducción castellana de Pedro Schwartz, *La Miseria del Historicismo*. Madrid, Taurus, 1961. Reeditada en Alianza Editorial, 1973.]

— *Conjectures and Refutations*, 1963, 1965, 1969. [Traducción castellana de Néstor Míguez, *El Desarrollo del conocimiento Científico. Conjeturas y Refutaciones*. Buenos Aires, Paidos, 1967.]

* — *Of Clouds and Clocks*, 1965. (Véase el capítulo 6.)

* — "Naturgesetze und theoretische Systeme" en *Gesetz und Wirklichkeit*, ed. Simon Moser (1948), 1949. (Traducido aquí en el Apéndice de este volumen.)

— "Die Zielsetzung der Erfahrungswissenschaft", en *Theorie und Realität*, ed. Hans Albert, 1964, págs. 73-86.

— "Quantum Mechanics without The Observer' ", en *Quantum Theory and Reality*, edición M. Bunge, 1967.

* — "Epistemology without a Knowing Subject", en *Logic, Methodology, and Philosophy of Science, 3*, eds. B. van Rootselaar y J. F. Staal, 1958, págs. 333-73. (Véase el capítulo 3.)

* — "On the Theory of Objective Mind", en *Akte des 14. Internationalen Kongresses für Philosophie. Wien*, 1968, *I*, págs. 25-33. [Hay traducción castellana de una versión de este artícuo, ligeramente aumentado en este volumen, de Víctor Sánchez de Zavala: "Sobre la Teoría de la Inteligencia Objetiva", en *Ensayos de Filosofía de la Ciencia*. (Simposio de Burgos. En torno a la obra de sir K. R. Popper.) Madrid, Tecnos, 1970, págs. 202-37.]

NOTA BIBLIOGRAFICA

En mis conferencias he expuesto repetidas veces la idea aquí discutida de que las teorías pueden *corregir* una ley "observacional" o "fenoménica" que pretenden explicar (como, por ejemplo, la tercera ley de Kepler). Una de esas conferencias sirvió de estímulo para corregir una ley fenoménica supuesta (véase el artículo de 1941, aludido en mi libro *Poverty of Historicism*, 1957, 1960, nota de la página 134 y sigs. [Traducción castellana, páginas 164-5]). Otra de esas conferencias fue publicada en el volumen de Simon Moser, *Gesetz und Wirklichkeit* (1948), 1949. Esta idea mía constituye también el "punto de partida" (como dice él en la pág. 92), de P. K. Feyerabend en su artículo "Explanation, Reduction and Empiricism" (en Herbert Feigl y Grover Maxwell, eds., *Minnesota Studies in the Philosophy of Science, 3*, 1962) que alude en [66] a este artículo (tal como se publicó por primera vez en *Ratio, I*, 1957). El reconocimiento de Feyerabend parece haber sido olvidado por los autores de diversos artículos sobre temas relacionados con éste.

6. SOBRE NUBES Y RELOJES*

APROXIMACIÓN AL PROBLEMA DE LA RACIONALIDAD Y LIBERTAD DEL HOMBRE

I

La persona que el año pasado pronunció en esta misma sala la primera conferencia en Memoria de Arthur Holly Compton era más afortunada que yo, pues llegó a conocerle personalmente, mientras que yo nunca he tenido esa suerte [1].

Sin embargo, comencé a oír hablar de él en mis tiempos de estudiante, allá por los años veinte, sobre todo a partir de 1925 cuando el famoso experimento de Compton y Simon [2] refutó la hermosa, aunque efímera, teoría de Bohr, Slater y Kramers [3]. Esta refutación fue uno de los acontecimientos decisivos de la historia de la teoría cuántica, puesto que de la crisis creada por ella surgió la llamada "nueva teoría cuántica" —las teorías de Born y Heisenberg, Schrödinger y Dirac.

Era la segunda vez que las contrastaciones experimentales de Compton habían desempeñado un papel crucial en la historia de la teoría cuántica. La vez primera fue, evidentemente, cuando se descubrió el efecto Compton, la primera contrastación independiente (como el propio Compton señaló) [4] de la teoría de los cuantos de luz o fotones de Einstein.

Años más tarde, durante la segunda guerra mundial, descubrí con sorpresa y satisfacción que Compton no sólo era un gran físico, sino también un filósofo genuino y audaz. También encontré que sus intereses y objetivos filosóficos coincidían con los míos en algunos puntos importantes. Lo descubrí, casi por casualidad, al dar con las extraordinarias Conferencias Terry

---

* Se trata de la segunda Conferencia en Memoria de Arthur Holly Compton, presentada en la Universidad de Washington el 21 de abril de 1965.

[1] Cuando llegué a Berkeley en febrero de 1962, procuré por todos los medios conocer a Compton, pero murió antes de que pudiésemos llegar a vernos.

[2] A. H. Compton y A. W. Simon, *Phys. Rev., 25,* 1925, págs. 309 y sigs. (Véase también W. Bothe y H. Geiger, *Zeit. f. Phys. 26,* 1924, págs. 44 y sigs.; y *32,* 1925, páginas 639 y sigs.; *Naturwissenschaften, 13,* 1925, pág. 440.)

[3] Bohr, H. A. Kramers y J. C. Slater, *Phil. Mag., 47,* 1924, págs. 785 y sigs. y *Zeitschr f. Phys., 24,* 1924, págs. 69 y sigs. Véase también A. H. Compton y S. K. Allison, *X-Rays in Theory and Experiment,* 1935; por ejemplo, las págs. 211-27.

[4] Véase el capítulo I, sección 19 del libro de Compton y Allison (citado en la nota 3).

que Compton había publicado en 1935 en un libro titulado *The Freedom of Man* [5].

Como ustedes habrán visto, he incorporado el título del libro de Compton, *The Freedom of Man,* al título que he dado a esta conferencia. Lo he hecho para poner de manifiesto que mi conferencia está íntimamente relacionada con su libro. Es más, trataré de discutir los mismos problemas de que se ocupaba Compton en los dos primeros capítulos de su libro, así como en el segundo capítulo de otro de sus libros, *The Human Meaning of Science* [6].

No obstante, para evitar malentendidos, he de subrayar que la conferencia de hoy no versa fundamentalmente acerca de los libros de Compton, sino que constituye más bien un intento de mirar con ojos nuevos los mismos viejos problemas filosóficos que atacaba en ambos libros, tratando de darles una nueva solución. Me parece que la solución esquemática y aproximativa que voy a bosquejar aquí encaja perfectamente con los objetivos fundamentales de Compton y espero —y además creo— que él estaría de acuerdo.

II

El propósito central de mi conferencia viene dado por el intento de plantearles a ustedes estos viejos problemas de un modo sencillo y tajante. Pero antes debo decir algo sobre las *nubes y relojes* que aparecen en el título de mi conferencia.

Mis nubes pretenden representar los sistemas físicos que, como los gases, son altamente irregulares, desordenados y más o menos impredictibles. Supondré que tenemos ante nosotros un esquema u ordenamiento a cuya izquierda está situada una nube muy perturbada o desordenada. En el otro extremo de nuestro ordenamiento, a la derecha, podemos colocar un reloj de péndulo, un reloj de precisión que intenta representar los sistemas físicos regulares, ordenados y de comportamiento altamente predictible.

De acuerdo con lo que se podría denominar la visión de las cosas según el sentido común, algunos fenómenos naturales, como el tiempo atmosférico o el paso de las nubes, son difíciles de predecir: hablamos de los "caprichos del tiempo". Por otro lado, hablamos de "precisión propia de un mecanismo de relojería" para describir fenómenos muy regulares y predictibles.

Entre ambos extremos —las nubes a la izquierda, los relojes a la derecha— podemos colocar muchísimas cosas, procesos y fenómenos naturales. Los cambios de estación son relojes un tanto imprecisos, por lo que tendríamos que situarlos hacia la derecha, aunque no demasiado. Supongo que será fácil ponerse de acuerdo en que los animales han de estar no muy

---

[5] A. H. Compton, *The Freedom of Man,* 1935 (3.ª ed., 1939). Este libro se basa fundamentalmente en las conferencias de la Fundación Terry, pronunciadas por Compton en Yale en 1931, así como en otra serie de conferencias pronunciadas después de las anteriores.

[6] A. H. Compton, *The Human Meaning of Science,* 1940.

alejados de las nubes, a la izquierda, mientras que las plantas habrán de colocarse algo más cerca de los relojes. Por lo que respecta a los animales, habrá que colocar a un joven cachorro algo más a la izquierda que a un perro viejo. Los coches también dispondrán de un lugar en alguna parte de nuestra ordenación de acuerdo con su precisión: imagino que un Cadillac estará bastante a la derecha, y mucho más un Rolls-Royce, que estará muy próximo al mejor de los relojes. Tal vez el *sistema solar* haya de situarse más a la derecha [7].

Como ejemplo típico e interesante de nube recurriré a una nube o enjambre de moscas o mosquitos. Los mosquitos individuales que, como las moléculas de un gas, forman todos juntos un enjambre se mueven de un modo asombrosamente irregular. Es casi imposible seguir el vuelo de un mosquito particular, aunque todos ellos sean lo suficientemente grandes como para ser visibles con claridad.

Aparte del hecho de que las velocidades de los mosquitos no muestran una dispersión muy amplia, los mosquitos nos suministran una imagen excelente de los movimientos irregulares de las moléculas en una nube de gas o de las diminutas gotas de agua de una nube tormentosa. Evidentemente, hay diferencias. El enjambre no se disuelve ni se difunde, sino que se mantiene perfectamente unido. Es algo sorprendente, si tenemos en cuenta el carácter desordenado del movimiento de los distintos mosquitos; pero es algo análogo a una nube de gas de tamaño considerable (como nuestra atmósfera o el sol), que se mantiene unida mediante fuerzas gravitatorias. En el caso de los mosquitos, el que se mantengan juntos puede explicarse fácilmente suponiendo que, aunque tengan un vuelo irregular en todas direcciones, aquellos que sienten que se alejan de la muchedumbre, tornan hacia su parte más densa.

Esta suposición explica de qué modo se mantiene unido el grupo aunque no tenga ni jefe ni estructura —no hay más que una distribución estadística aleatoria que surge del hecho de que cada mosquito hace exactamente lo que quiere de un modo anárquico y aleatorio unido al hecho de que no quiere separarse demasiado de sus compañeros.

Tal vez algún mosquito filosófico pueda decir que la sociedad de los mosquitos es una gran sociedad o, cuando menos, una buena sociedad, puesto que es la más igualitaria, libre y democrática de las sociedades imaginables.

Sin embargo, como autor de un libro sobre *La Sociedad Abierta*, tengo que decir que la de los mosquitos no es una sociedad abierta, pues creo que una de las características de este tipo de sociedad consiste en valorar positivamente, no sólo la forma democrática de gobierno, sino también la libertad de asociación, junto con la protección y fomento de la formación de sub-sociedades con diversas opiniones y creencias. Ahora bien, todo mos-

---

[7] Véanse las notas 11 y 19, más adelante, para las imperfecciones del sistema solar.

quito razonable habrá de admitir que en su sociedad falta este tipo de pluralismo.

No obstante, no pretendo discutir hoy ninguna cuestión social o política relacionada con el problema de la libertad. Además, no pretendía utilizar la nube de mosquitos a modo de ejemplo de sistema *social*, sino más bien como ilustración de un sistema *físico* nebuloso; es decir, como ejemplo o paradigma de nube desordenada o muy irregular.

Del mismo modo que ocurre con muchos sistemas físicos, biológicos y sociales, el enjambre de mosquitos puede describirse como un "todo". Mi conjetura de que se mantienen unidos por una especie de atracción que ejerce la parte más densa sobre los mosquitos individuales que se alejan demasiado de la muchedumbre, muestra que incluso dicho todo ejerce una especie de acción o control sobre sus elementos o partes. Sin embargo, este "todo" puede aducirse en contra de la creencia "holística" tan extendida, según la cual el "todo" es *siempre* más que la mera suma de sus partes. No niego que algunas veces pueda ser así [8]. De todos modos, el enjambre de mosquitos suministra el ejemplo de un todo que no es más que la suma de sus partes (y en un sentido muy preciso), ya que no sólo se puede describir completamente mediante los movimientos de todos los mosquitos individuales, sino que además el movimiento del todo es, en este caso, exactamente la suma (vectorial) de los movimientos de sus miembros constituyentes dividida por el número de ellos.

Un ejemplo (semejante en muchos sentidos) de sistema biológico que constituye un "todo", con un control sobre los movimientos en gran medida irregulares de sus partes, viene dado por una familia que va de excursión al campo —los padres con unos niños y un perro— y vagabundea durante horas por el bosque sin alejarse nunca mucho del coche familiar (que actúa, como si dijéramos, a modo de centro de atracción). Puede decirse que este sistema es aún más nebuloso que nuestra nube de mosquitos; es decir, menos regular respecto al movimiento de sus partes.

Espero que se hayan hecho ustedes una idea de mis dos prototipos o paradigmas (las nubes a la izquierda y los relojes a la derecha), así como del modo en que podemos disponer entre ellos diversos tipos de cosas y sistemas. Estoy seguro de que ustedes se han hecho una idea vaga y general de dicha ordenación y no tienen por qué preocuparse si esa idea es aún un tanto brumosa o nebulosa.

---

[8] Véase la sección 23 de mi libro *The Poverty of Historicism*, 1957 y ediciones sucesivas [Traducción castellana de Pedro Schwartz, *La Miseria del Historicismo*, Madrid, Taurus, 1961. Reeditado en Alianza Editorial, Madrid, 1973] en la que critico el criterio "holístico" de "todo" (o "*Gestalt*"), mostrando que dicho criterio ("un todo es más que la suma de sus partes") lo satisfacen incluso los ejemplos típicos que ponen los partidarios del holismo para mostrar lo que es un todo, como por ejemplo, un simple montón de piedras. (Nótese que no niego que haya todos, me limito a plantear objeciones a la superficialidad de las teorías "holísticas".)

III

La disposición que he descrito es a primera vista totalmente aceptable para el sentido común; además, en nuestros días, incluso es aceptable en física. No obstante, no era ese el caso en los doscientos cincuenta últimos años; la revolución newtoniana, una de las revoluciones más grandes de la historia, conducía al rechazo de la ordenación de sentido común que he tratado de presentarles a ustedes. Esto es así porque, como todo el mundo * piensa, una de las cosas que la revolución newtoniana ha demostrado es la siguiente proposición asombrosa:

Toda las nubes son relojes, incluso las nubes más nebulosas.

Esta proposición, "Todas las nubes son relojes", puede considerarse una formulación abreviada del punto de vista que denominaré "*determinismo físico*".

El determinismo físico, además de decir que todas las nubes son relojes, afirmará que nuestras ordenaciones de sentido común —con las nubes a la izquierda y los relojes a la derecha— son erróneas porque todo debiera situarse en el extremo de la derecha. Dirá que con todo nuestro sentido común hemos situado las cosas, *no de acuerdo con su naturaleza, sino, simplemente, de acuerdo con nuestra ignorancia.* Nuestra ordenación, nos dirá, no hace más que reflejar el hecho de que conocemos con cierto detalle el funcionamiento de las partes de un reloj, o del sistema solar, mientras que no tenemos la menor idea acerca de cómo interrelacionan en *detalle* las partículas que componen una nube de gas o un organismo. También dirá que, una vez conseguido dicho conocimiento, veremos que esa nube de gas o ese organismo tienen una estructura de relojería como la que pueda tener nuestro sistema solar.

Evidentemente, la teoría de Newton no indicaba a los físicos que esto fuese así. De hecho, no se ocupaba en absoluto de las nubes, sino de las balas de cañón, de las mareas y fundamentalmente de los planetas cuyos movimientos explicaba atribuyéndolos a leyes naturales muy simples. Sin embargo, su inmenso éxito en estos terrenos trastornó a los físicos, no sin razón.

Antes de la época de Newton y su predecesor, Kepler, los movimientos de los planetas se habían resistido a dejarse explicar e incluso describir plenamente. Aunque, como es natural, participaban en alguna medida del movimiento general e invariable del sistema rígido de las estrellas fijas, con todo se desviaban del movimiento del sistema casi del mismo modo que los mosquitos individuales se desvían del movimiento general del enjambre. Por tanto, los planetas, que no se diferencian mucho de las cosas vivas, parecían estar en una posición intermedia entre las nubes y los relojes. Sin embargo, los éxitos de Kepler y, mucho más aún, los de la

* Newton no sacó él mismo de sus teorías estas consecuencias "deterministas": véase más adelante las notas 11 y 16.

teoría newtoniana dieron la razón a aquellos pensadores que habían sospechado que los planetas eran de hecho relojes perfectos, pues sus movimientos se podrían predecir exactamente mediante la teoría newtoniana (se podían predecir en todos aquellos aspectos que habían defraudado anteriormente a los astrónomos por su aparente irregularidad).

La teoría newtoniana fue la primera teoría científica que tuvo un éxito real en la historia humana y, por cierto, un éxito tremendo. Constituía un conocimiento real, un conocimiento que superaba los sueños desaforados de las mentes más audaces. Constituía una teoría que explicaba con precisión, no sólo los movimientos de *todos* los astros en su tránsito por los cielos, sino además, los movimientos de los cuerpos terrestres como la caída de una manzana, los proyectiles y los relojes de péndulo. Y además, explicaba las mareas.

Todos los hombres de mentalidad abierta —los deseosos de aprender que se interesaban por el aumento del conocimiento— se convirtieron a la nueva teoría. La mayoría de los hombres de mentalidad abierta, en especial casi todos los científicos, creían que acabaría explicándolo todo, incluyendo, no sólo la electricidad y el magnetismo, sino también las nubes e incluso los organismos vivos. De este modo, el determinismo físico —la doctrina de que todas las nubes son relojes— se impuso como fe indiscutible entre las personas ilustradas; quien no abrazaba esta fe era tenido por oscurantistas o reaccionario [10].

IV

Charles Sanders Peirce, el gran matemático y físico americano que considero uno de los mayores filósofos de todos los tiempos, se encontraba en el pequeño grupo de los disidentes [11]. No ponía en cuestión la teoría de Newton; ya en el año 1892 mostró que aunque la teoría fuese verdadera no nos suministraría ninguna razón válida para creer que las nubes sean relojes perfectos. Aunque compartía con todos los físicos de su tiempo la creencia de que el mundo era un reloj que funcionaba de acuerdo con las leyes de Newton, no aceptaba la creencia de que dicho reloj, o cualquier otro, fuese

---

[10] La convicción de que el determinismo constituía una parte esencial de toda actitud racional o científica era generalmente aceptada, incluso por algunos de los principales opositores del "materialismo" (tales como Spinoza, Leibniz, Kant y Schopenhauer). Un dogma semejante, que también formaba parte de la tradición racionalista, era la creencia de que todo conocimiento comienza con la *observación* y procede por inducción. Cf. mis consideraciones acerca de ambos dogmas en mi libro *Conjectures and Refutations*, 1963, 1965, 1969, 1972, págs. 122 y sigs. [Traducción castellana de Néstor Míguez, *El Desarrollo del Conocimiento Científico. Conjeturas y Refutaciones*, Buenos Aires, Paidos, 1967; págs. 144 y sigs.]

[11] El propio Newton puede contarse entre los escasos disidentes, pues consideraba que incluso el sistema solar era *imperfecto* y, consiguientemente, susceptible de aniquilación. A causa de esta opinión fue acusado de impiedad, de "poner en duda la sabiduría del autor de la naturaleza" (como nos dice Henry Pemberton en su *A View of Sir Isaac Newton's Philosophy*, 1728, pág. 180).

perfecto hasta en los menores detalles. Señalaba que, en cualquier caso, partiendo de la experiencia no es posible pretender alcanzar el conocimiento de algo semejante a un reloj perfecto; ni siquiera de algo que se asemeje débilmente a la absoluta perfección supuesta por el determinismo físico. Podría citar uno de los brillantes comentarios de Peirce: "...quien está entre bastidores" (Peirce habla aquí como experimentador) "...sabe que las comparaciones más refinadas [incluso] de las masas [y] longitudes..., que exceden con mucho en precisión a todas las demás mediciones [físicas], ...están muy por debajo del grado de precisión de los cómputos bancarios y que... la determinación de las constantes físicas... está aproximadamente a la par que las mediciones de alfombras y cortinas de un tapicero..." [12]. De ahí concluía Peirce que tenemos plena libertad para suponer la existencia de una cierta *holgura o imperfección* en todos los relojes, lo cual permitía introducir un *elemento de azar*. Por tanto, Peirce conjeturaba que el mundo no sólo no estaba *estrictamente* regulado por las *leyes newtonianas*, sino que, a la vez, también estaba dominado por *leyes de azar*, aleatorias o de desorden: por leyes de *probabilidad* estadística. El mundo se convertía así en un sistema trabado de nubes y relojes de tal modo que incluso los mejores relojes mostrarían cierto grado de nubosidad *en su estructura molecular*. Que yo sepa, Peirce fue el primer físico y filósofo postnewtoniano que tuvo la osadía de sostener la opinión de que *todos los relojes son nubes* hasta cierto punto; o, en otras palabras, que *sólo existen nubes*, aunque éstas pueden poseer diversos grados de nubosidad.

Peirce defiende su tesis señalando, sin duda correctamente, que todos los cuerpos físicos, incluso los diamantes de un reloj, están dotados de movimiento molecular debido al calor [13], movimiento similar al de las moléculas de un gas o al de los mosquitos individuales de un enjambre.

Los contemporáneos de Peirce recibieron sus ideas con escaso interés.

---

[12] *Collected Papers of Charles Sanders Peirce, 6*, 1935, 6.44, pág. 35. Evidentemente, puede que otros físicos hayan desarrollado ideas similares; pero, a parte de Newton y Peirce, sólo conozco a uno: el Profesor Franz Exner de Viena. Su discípulo Schrödinger habla acerca de las opiniones de Exner en su libro *Science, Theory and Man*, 1957, págs. 71, 133, 142 y sigs. (Este libro se había publicado anteriormente con el título *Science and the Human Temperament*, 1935, al que alude Compton en *The Freedom of Man*, pág. 29.) Cf. también la nota 26 más adelante.

[13] C. S. Peirce, *Op. cit., 6*, 6.47, pág. 37 (publicada por primera vez en 1892). El pasaje, a pesar de su brevedad, es de lo más interesante porque anticipa (nótese la consideración acerca de las fluctuaciones en las mezclas explosivas) algunas discusiones sobre macro-objetos derivados de la amplificación de las indeterminaciones de Heisenberg. Al parecer, esta discusión comienza con un artículo de Ralph Lillie, *Science, 46*, 1927, págs. 189 y sigs., al que alude Compton en *The Freedom of Man*, página 50. Desempeña una función considerable en el libro de Compton, págs. 48 y siguientes. (Nótese que Compton pronunció sus conferencias Terry en 1931). La *op. cit.* de Compton, nota 3, págs. 51 y sigs., contiene una interesante comparación cuantitativa de los efectos aleatorios debidos al movimiento molecular del calor (la indeterminación en que pensaba Peirce) con la indeterminación de Heisenberg. La discusión fue desarrollada por Bohr, Pascual Jordan, Fritz Medicus, Ludwig von Bertalanffy y muchos otros, especialmente, en época reciente, Walter Elsasser, *The Physical Foudations of Biology*, 1958.

Aparentemente sólo un filósofo las tomó en cuenta para atacarlas [14]. Los físicos parecen haberlas ignorado e, incluso hoy, la mayoría de ellos creen que si aceptásemos que la mecánica clásica newtoniana es verdadera, nos veríamos obligados a aceptar el determinismo físico junto con la proposición según la cual todas las nubes son relojes. Los físicos sólo estuvieron dispuestos a abandonar el determinismo físico tras la caída de la física clásica y el surgimiento de la nueva teoría cuántica.

Entonces cambiaron las tornas. El indeterminismo, que hasta 1927 era igual a oscurantismo, se convirtió en una moda generalizada hasta el punto de que algunos grandes científicos como Max Planck, Erwin Schrödinger y Albert Einstein fueron tenidos por viejos carcamales [15], apesar de estar a la vanguardia del desarrollo de la teoría cuántica. En cierta ocasión, oí cómo un físico joven y brillante motejaba de "antediluviano" a Einstein que aún vivía y seguía trabajando firmemente. Se pensaba que el diluvio que había tragado a Einstein era la nueva teoría cuántica surgida entre los años 1925-7, cuyo advenimiento fue posible gracias a las contribuciones de Einstein, contribuciones con las que sólo se pueden comparar a lo sumo las de otras siete personas.

V

Tal vez deba detenerme un momento aquí para exponer mi propia opinión acerca de esta situación y de las modas científicas. Creo que Peirce estaba en lo cierto cuando sostenía que en gran medida todos los relojes son nubes —incluso el más preciso de los relojes—. Pienso que se trataba de la inversión más importante de la tesis equivocada, según la cual todas las

[14] Me refiero a Paul Carus, *The Monist, 2*, 1892, págs. 560 y sigs., y 3, 1892, páginas 68 y sigs.; Peirce respondió en *The Monist, 3*, 1893, págs. 526 y sigs. (Véanse sus *Collected Papers, 6*, Apéndice A, págs. 390 y sigs.)

[15] La transformación total y repentina de la situación problemática puede calibrarse por el hecho de que a muchos de nosotros, viejos carcamales, nos parece que no hace tanto tiempo que los filósofos empiristas (véase, por ejemplo, Moritz Schlick, *Allgemeine Erkenntnislehre*, 2.ª ed., 1925, pág. 277) eran partidarios del determinismo físico. Sin embargo, hoy día, un defensor tan dotado e inspirado de Schlick, como es P. H. Nowell-Smith, desecha el determinismo por considerarlo un "fantasma dieciochesco" (*Mind, 63*, 1954, pág. 331; véase también la nota 37, más adelante). El tiempo pasa y no hay duda de que con el tiempo se resolverán todos nuestros problemas, sean o no fantasmas. Sin embargo, nosotros, los viejos carcamales, parecemos recordar de un modo bastante singular los tiempos de Plank, Einstein y Schlick y nos sentimos turbados cuando intentamos convencer a nuestras mentes perplejas y empantanadas de que estos grandes pensadores deterministas crearon sus fantasmas en el siglo dieciocho junto con Laplace, el creador del fantasma más famoso de todos (la "inteligencia sobrehumana" de su *Ensayo* de 1819, llamado comúnmente el "demonio de Laplace"; cf. Compton, *The Freedom of Man*, pág. 5 y sigs., y *The Human y Meaning of Science*, pág. 34, así como Alexander, citado más adelante en la nota 36). Sin embargo, puede que un esfuerzo aún mayor nos traiga a la memoria —nuestra falible memoria— un fantasma dieciochesco similar creado por un tal Carus (no el pensador decimonónico P. Carus, al que aludía en la nota 14, sino T. L. Carus, el autor de *Lucretius de Rerum Naturae*, II. 251-60, citado por Compton en *The Freedom of Man*, pág. 1).

nubes son relojes. Creo, además, que Peirce estaba en lo cierto al sostener que su opinión era compatible con la física clásica de Newton [16]. Creo que su punto de vista es compatible de un modo aún más claro con la teoría de la relatividad (especial) de Einstein y, aún más, con la nueva teoría cuántica. En otras palabras, soy indeterminista —como Peirce, Compton y la mayoría de los físicos contemporáneos—. Además creo, con casi todo el mundo, que Einstein cometió el error de tratar de mantener el determinismo a toda costa. (Tal vez deba decir que discutí con él estas cuestiones y no me pareció una persona inflexible.) Sin embargo, creo que también están profundamente equivocados aquellos físicos modernos que desprecian la crítica de Einstein a la teoría cuántica por considerarla antediluviana. Nadie puede menos de admirar la teoría cuántica, cosa que hacía Einstein de todo corazón; pero su crítica a la interpretación en boga de la teoría —la interpretación de Copenhague— así como las críticas de De Boglie, Schrödinger, Bohm, Vigier y últimamente Landé han sido dejadas de lado con demasiada ligereza por parte de la mayoría de los físicos [17]. En la ciencia hay modas y algunos científicos se suben al carro triunfal con la misma celeridad que algunos pintores y músicos. Mas, aunque las modas y los carros triunfales atraigan a los débiles, no hay que fomentarlos, sino que es preciso resistirse [18]. Por esta razón es siempre valiosa una crítica como la de Einstein: siempre se puede aprender algo de ella.

---

[16] Desarrollé esta opinión en 1950 en un artículo titulado "Indeterminism in Quantum Physics and in Classical Physics", *British Journal for the Philosophy of Sciencie, 1,* 1950, núm. 2, págs. 117-33 y núm. 3, págs. 173-95. Desgraciadamente, cuando lo escribí no tenía ni idea de las opiniones de Peirce (véanse las notas 12 y 13). Tal vez deba mencionar aquí que de este artículo mío he sacado la idea de oponer las *nubes* a los *relojes*. Desde 1950, cuando se publicó aquel artículo, ha tomado un gran impulso la discusión de los elementos indeterministas de la física clásica. Véase Leon Brillouin, *Scientific Uncertainty and Información*, 1964 (libro con el que no estoy, en absoluto, plenamente de acuerdo), así como las referencias bibliográficas que aparecen allí, concretamente, en las páginas 38, 105. 127, 151 y sigs. A dichas referencias habría que añadir en particular el excelente artículo de Jacques Hadamard sobre las líneas geodésicas en las superficies "corniformes" de cuvartura negativa, *Journal de Mathématiques pures et appliquées,* 5.ª serie, 4, 1898, págs. 27 y sigs.

[17] Véase también mi libro *The Logic of Scientific Discovery* [traducción castellana de Víctor Sánchez de Zavala: *La Lógica de la Investigación Científica,* Madrid, Tecnos, 1962], especialmente el nuevo Apéndice * XI y el capítulo IX que contiene una crítica válida en líneas generales, aunque en vista de la crítica de Einstein al Apéndice * XII hube de trazar el experimento mental (de 1934) descrito en la sección 77. No obstante, dicho experimento puede sustituirse por el famoso experimento mental de Einstein, Podolsky y Rosen discutido en los Apéndices XXI y * XII. Véase también mi artículo "The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory", en *Observation and Interpretation,* ed. por S. Körner, 1957, págs. 65-70 y 83-9.

[18] Esta frase fue pensada para criticar algunos de los puntos de vista contenidos en el interesante y estimulante libro de Thomas S. Kuhn. *The Structure of Scientific Revolutions,* 1963. [Traducción castellana de Agustín Contín. *La Estructura de las Revoluciones Científicas,* México, Fondo de Cultura Económica, 1972; la fecha de publicación original es 1962.]

VI

Arthur Holly Compton fue uno de los primeros en dar la bienvenida a la nueva teoría cuántica y al nuevo indeterminismo físico de Heisenberg de 1927. Compton invitó a Heisenberg a dar un ciclo de conferencias en Chicago que tuvo lugar en la primavera de 1929. Este ciclo fue la primera exposición completa que hizo Heisenberg de su teoría, siendo publicadas sus conferencias en su primer libro publicado al año siguiente en la University of Chicago Press con un prólogo de Compton [19]. En dicho prólogo, Compton daba la bienvenida a la nueva teoría cuyo advenimiento habían favorecido sus experimentos al refutar a su antecesora inmediata [20]. Sin embargo, también daba un toque de atención, anticipándose a unas consideraciones muy semejantes de Einstein, en el sentido de que no debíamos considerar "completa" la nueva teoría cuántica [21] —"este capítulo de la historia de la física", como la denominaba Compton sabia y generosamente. Aunque Bohr rechazase este punto de vista, hemos de recordar que la nueva teoría fallaba, por ejemplo, a la hora de suministrar siquiera fuese una pista sobre el neutrón, descubierto por Chadwick aproximadamente un año más tarde y que no fue más que el comienzo de una larga serie de nuevas partículas elementales cuya existencia no había sido prevista por la nueva teoría cuántica (si bien es cierto que la teoría de Dirac podía haber dado pie a la deducción de la existencia del positrón) [22].

En el mismo año 1931, Compton, en sus conferencias de la Fundación Terry [23], fue uno de los primeros en analizar las implicaciones humanas y, en general, biológicas [24] del nuevo indeterminismo físico, poniendo de ma-

---

[19] Véase Werner Heisenberg, *The Physical Principles of the Quantum Theory*,, 1930.

[20] Me refiero a la refutación de Compton de la teoría de Bohr, Slater y Kramers; véase la nota 3, *supra;* véase también la propia alusión de Compton en *The Freedom of Man*, pág. 7 (última frase) y en *The Human Meaning of Science*, página 36.

[21] Cf. el prólogo de Compton al libro de Heisenberg, *op. cit.*, págs. III y sigs., así como sus consideraciones acerca del *carácter incompleto* de la mecánica cuántica en *The Freedom of Man*, pág. 45 (con una alusión a Einstein), y en *The Human Meaning of Science*, pág. 42. Compton aprobaba el carácter incompleto de la mecánica cuántica, oponiéndose a Einstein que veía en él una debilidad de la teoría. En respuesta a Einstein, Niels Bohr dijo (como había dicho antes J. von Neumann) que la teoría era *completa* (tal vez en otro sentido del término). Véase, por ejemplo, A. Einstein, B. Podolsky y N. Rosen, *Physical Review. 42*, 1935, págs. 777-80, así como la respuesta de Bohr en *48*, 1935, págs. 1.935 y sigs. Véase también A. Einstein, *Dialectica, 2*, 1948, págs. 320-4 y Bohr, págs. 312-9 del mismo volumen. Más sobre la discusión anterior entre Einstein y Niels Bohr: P. A. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist*, 1949, págs. 201-41 y especialmente, 668-74, así como la carta de Einstein publicada en mi libro *The Logic of Scientific Discovery*, págs. 457-64 [trad. cit., páginas 426-32] y también págs. 445-56 [ibid., 414-25].

[22] Véase la historia de este descubrimiento, tal como la cuenta N. R. Hanson, *The Concept of the Positron*, 1963, capítulo IX.

[23] A. H. Compton, *The Freedom*... (Cit. supra, nota 5).

[24] Véanse especialmente los pasajes dedicados a la "evolución emergente" en *The Freedom of Man*, págs. 90 y sigs.; cf. *The Human Meaning of Science*, págs. 73.

nifiesto por qué había dado la bienvenida con tanto entusiasmo a la nueva teoría: para él resolvía no sólo problemas físicos, sino también problemas biológicos y filosóficos, especialmente estos últimos, algunos de ellos relacionados con la ética.

VII

Para demostrarlo, citaré el sorprendente pasaje inicial del libro de Compton, *The Freedom of Man*:

> La cuestión fundamental de la moralidad, que constituye no sólo un problema religioso vital, sino también un tema de activa investigación científica es: ¿Es el hombre un agente libre?
>
> Si... los átomos que componen nuestros cuerpos obedecen leyes físicas tan inmutables como los movimientos de los planetas, ¿por qué esforzarse?, ¿qué puede importar la magnitud de nuestro esfuerzo, si nuestros actos ya están predeterminados por leyes mecánicas...? [25].

Compton describe aquí lo que denominaré "*la pesadilla del determinista físico*". Un mecanismo de relojería físicamente determinado es, ante todo, completamente independiente: en un mundo físico totalmente determinado no hay lugar sencillamente para una intervención externa. Todo lo que ocurre en este mundo está predeterminado físicamente, incluyendo nuestros movimientos y, por tanto, nuestros actos. Así pues, todos nuestros pensamientos, sentimientos y esfuerzos no pueden tener ninguna influencia práctica sobre lo que ocurre en el mundo físico: son, si no meras ilusiones, a lo sumo subproductos superfluos ("epifenómenos") de acontecimientos físicos.

De este modo, las ensoñaciones del físico newtoniano, que quería demostrar que todas la nubes son relojes, amenazan volverse pesadillas y el esfuerzo por ignorar esta situación ha conducido a una especie de personalidad intelectual escindida. Pienso que Compton estaba agradecido a la nueva teoría cuántica por haberle salvado de esta difícil situación intelectual. Así pues, escribe en *The Freedom of Man*: "Los físicos rara vez... se sienten molestos por el hecho de que si... se aplican al hombre... leyes... totalmente deterministas, ...él mismo es un autómata" [26]. También en *The Human Meaning of Science* pone esto mismo de relieve:

> Por lo que respecta a lo que yo pienso acerca de este tema vital, me encuentro en una situación mucho más satisfactoria que la que hubiese podido alcanzar en un estadio anterior de la ciencia. Si suponemos que los enunciados de las leyes físicas

[25] Cf. *The Freedom of Man*, pág. 1.

[26] Cf. *The Freedom of Man*, pág. 26 y sig.; véanse también las págs. 27 y siguientes (el último párrafo que comienza en la pág. 27). Tal vez deba recordar al lector que mis puntos de vista difieren un tanto del pasaje citado, ya que, como Peirce, considero lógicamente posible que las *leyes* de un sistema sean newtonianas (y de ese modo, deterministas *prima facie*) y que a pesar de todo el sistema siga siendo inde-

son correctos, hemos de aceptar (como hicieron la mayoría de los filósofos) el carácter ilusorio del sentimiento de libertad o, en caso de que consideráse mos efectivo el sentimiento de [libre] elección, entonces habríamos de concluir que el enunciado de las leyes físicas no es digno de confianza. El dilema ha sido una cosa incómoda... [27].

Más adelante, en el mismo libro, Compton resume vigorosamente la situación con las palabras: "...ya no tiene ninguna justificación emplear las leyes físicas como testimonios en contra de la libertad humana" [28].

Estas citas de Compton muestran, con toda claridad, que se había sentido incómodo, antes que Heisenberg, con lo que llamo aquí la pesadilla del determinista físico y que había tratado de escapar de dicha pesadilla, adoptando algo así como una escisión intelectual de la personalidad. O, como él mismo dice: "Nosotros [los físicos] hemos preferido sencillamente no prestar atención a las dificultades..." [29]. Compton dio la bienvenida a la nueva teoría que le libraba de todo esto.

Creo que el único aspecto del problema del determinismo que vale la pena discutir seriamente es, precisamente, el problema que le preocupaba a Compton: el problema surgido de una teoría física que describe el mundo como un sistema *físicamente completo* o *físicamente cerrado* [30]. Por sistema físicamente cerrado entiendo un conjunto o sistema de entidades físicas como átomos, partículas elementales, fuerzas físicas o campos de fuerzas que interactúan entre sí —y *solo* entre sí— de acuerdo con leyes definidas de interacción que no dejan lugar a interacciones o interferencias con nada externo a ese conjunto o sistema cerrado de entidades físicas. Es esta "clausura" del sistema lo que crea la pesadilla determinista [31].

---

terminista, pues el sistema a que se aplican las leyes puede ser intrínsecamente impreciso en el sentido, por ejemplo, de que no se puede decir que sus coordenadas sean números racionales (más bien que irracionales). La siguiente consideración (véase Schrödinger, *op. cit.*, pág. 143) es también muy relevante: "...el teorema energía-momento sólo nos suministra *cuatro* ecuaciones, dejando de ese modo indeterminado en gran medida el proceso elemental, a pesar de que concuerde con ellas". Véase también la nota 16.

[27] Cf. *The Human Meaning of Science*, pág. IX.

[28] Ibid., pág. 42.

[29] Cf. *The Freedom of Man*, pág. 27.

[30] Supongamos que nuestro mundo físico es un sistema *físicamente cerrado* que contiene elementos de azar. Obviamente, no será determinista; pero a pesar de todo, los propósitos, ideas, deseos y esperanzas no tendrán ninguna influencia sobre los acontecimientos físicos de este mundo; suponiendo que existan, serán totalmente redundantes: serán lo que se suele llamar "epifenómenos". (Nótese que un sistema físico determinista tiene que ser cerrado, si bien un sistema cerrado puede ser indeterminista. Así pues, "el indeterminismo no basta", como explicaremos en la sección X más adelante; véase también la nota 42.

[31] Kant se sintió profundamente atormentado por esta pesadilla de la que no consiguió escapar; véase el excelente planteamiento que hace Compton de "la vía de escape kantiana" en su libro *The Freedom of Man*, págs. 67 y sig. (En la línea 2 de la pág. 68 habría que tachar las palabras "*de la Razón Pura*".) Quizá deba decir aquí que no estoy de acuerdo con todas las cosas que Compton nos dice en el terreno de la filosofía de la ciencia. He aquí algunos ejemplos de cosas que no comparto: la aprobación de Compton del positivismo o fenomenismo de Heisenberg (*The Freedom of Man*, pág. 31) y ciertas consideraciones (en *op. cit.*, nota 7 de la pág. 20) que Compton atribuye a Carl Eckart: aunque Newton mismo no era, al parecer, determinista (cf. nota 11) no creo que se pueda discutir la idea bien precisa de *determinismo físi-*

VIII

Quisiera hacer ahora una breve digresión para comparar el problema del determinismo físico, que considero de fundamental importancia, con el problema nada serio por el que lo han sustituido muchos filósofos y psicólogos siguiendo a Hume.

Hume interpretaba el determinismo (que denominaba "la doctrina de la necesidad" o "doctrina de la conjunción constante") en el sentido de la doctrina según la cual "causas semejantes producen siempre efectos semejantes" y "los efectos semejantes se siguen necesariamente de causas semejantes" [32]. Sostenía, por lo que respecta concretamente a las acciones y voliciones humanas, que "un espectador puede normalmente inferir nuestras acciones de nuestros motivos y carácter y, aún cuando él no pueda hacerlo, concluye en general que sería posible si conociese perfectamente las circunstancias de nuestra situación y temperamento y los más secretos principios de nuestra... disposición. Ahora bien, ésta es la verdadera esencia de la necesidad..." [33]. Los seguidores de Hume lo formularon del siguiente modo: nuestras acciones, voliciones, gustos o preferencias son "causadas" psicológicamente por experiencias anteriores ("motivos") y, en última instancia, por nuestra herencia y por el medio.

Mas esta doctrina que denominaré determinismo *filosófico* o *psicológico*, no sólo es algo muy diferente del determinismo *físico*, sino que además es algo que difícilmente puede tomar en serio un determinista físico versado en la materia. La razón estriba en que la tesis del determinismo filosófico, según la cual "efectos semejantes poseen causas semejantes" o "todo acontecimiento tiene una causa", es tan vaga que resulta perfectamente compatible con el indeterminismo físico.

Lo único que afirma el *indeterminismo* —o, más exactamente, el indeterminismo físico— es que *no todos* los acontecimientos del mundo físico están predeterminados con absoluta precisión en todos sus detalles infinitesimales. Al margen de esto, resulta compatible prácticamente con cualquier grado de regularidad que se desee y, por tanto, no implica la opinión de que haya "acontecimientos" sin "causas", por la sencilla razón de que las palabras "acontecimientos" y "causas" son lo suficientemente vagas como para permitir la compatibilidad del indeterminismo físico con la doctrina de que todo acontecimiento tiene una causa. Frente al determinismo físico, que exige una predeterminación física completa e infinitamente precisa sin excepción *alguna*, el indeterminismo físico sólo afirma que el determi-

---

co en términos de una vaga "ley de causalidad"; tampoco estoy de acuerdo con que Newton fuese un fenomenista en un sentido similar a aquel en que se decía, en los años treinta, que Heisenberg era fenomenista (o positivista).

[32] David Hume, *A Treatise of Human Nature*, 1739 (*ed. L. A. Selby-Bigge*, 1888 y reediciones, pág. 174 [traducción castellana de Vicente Viqueira, *Tratado de la Naturaleza Humana*, Madrid, Calpe, 1923, pág. 278]; véase además, por ejemplo, las págs. 173 y 87 [págs. 277 y sigs. y 143].

[33] Hume, *op. cit.*, pág. 408 y sigt. [trad cit., pág. 199].

nismo es falso y que *al menos* hay, aquí y allá, algunas excepciones a la determinación precisa.

Por tanto, incluso la fórmula "Todo acontecimiento *físico* observable o medible posee una causa física observable o mediable" sigue siendo compatible con el indeterminismo físico, por la sencilla razón de que ninguna medida puede ser infinitamente precisa. La cuestión fundamental del determinismo físico, que se basa en la dinámica newtoniana, es que afirma la existencia de un mundo dotado de una precisión matemática absoluta y, aunque al hacerlo trascienda el dominio de la observación posible, como vio Peirce, no obstante es contrastable en principio con un grado de precisión aceptable; y de hecho resistió contrastaciones sorprendentemente precisas.

Por el contrario, la fórmula "Todo acontecimiento posee una causa" no dice nada preciso; concretamente, si consideramos las leyes de la psicología, vemos que no hay en ella el menor atisbo de precisión. Esta es la situación no sólo de la psicología "introspectiva" o "mentalista", sino también de la psicología "conductista". Es bastante obvio que ocurre así en la psicología mentalista, pero también los conductistas predicen, *en el mejor de los casos,* que, en ciertas condiciones, la rata tardará veinte o veintidós segundos en recorrer un laberinto: no tiene la menor idea de cómo hacer predicciones cada vez más precisa —*en principio, precisas sin limitaciones*— especificando cada vez más las condiciones experimentales. La causa de ello es que las "leyes" conductistas no son ecuaciones diferenciales como las de la física newtoniana; además, todo intento de introducir dichas ecuaciones nos hará abandonar el campo de la psicología para hacer fisiología y, en última instancia, física. Por tanto, nos veremos llevados hacia el problema del *determinismo físico.*

Como hizo notar Laplace, el determinismo físico implica que todo acontecimiento físico perteneciente al futuro lejano (o al pasado remoto) es predictible (o retrodictible) con un grado aceptable de precisión, a condición de que conozcamos de modo suficiente el estado actual del mundo físico. Por otro lado, la tesis del determinismo filosófico (o psicológico) del tipo de Hume, incluso en su versión más fuerte, se limita a decir que cualquier diferencia *observable* entre dos acontecimientos se relaciona mediante una ley, quizá aún desconocida, con determinada diferencia —tal vez observable— en el estado anterior del mundo. Obviamente, se trata de una afirmación mucho más débil que, dicho sea de paso, podemos seguir manteniendo aún cuando la mayoría de nuestros experimentos, realizados en condiciones *aparentemente* "totalmente iguales", arrojen resultados diferentes. El propio Hume formuló estas cosas con toda claridad. "Aún cuando los experimentos contrarios sean totalmente iguales", escribe, "no omitimos la noción de causa y necesidad, sino que... concluimos que el azar [aparente] ...reside tan sólo en... nuestro conocimiento imperfecto, no en las cosas mismas, que son en todo caso igualmente necesarias [i.e., determinadas], aunque en apariencia no son igualmente constantes o ciertas" [34].

[34] Hume, *op. cit.*, págs. 403 y sigs. [trad. cit., págs. 191-2]. Es interesante com-

Por eso el determinismo filosófico de Hume y, concretamente, el determinismo psicológico, carece de la garra del determinismo físico. Para la física newtoniana parecía realmente como si la aparente imprecisión de un sistema no se debiese más que a nuestra ignorancia, de manera que si tuviésemos una información completa acerca del sistema, desaparecería todo rastro de imprecisión. Por el contrario, la psicología nunca ha tenido este carácter.

Retrospectivamente, podemos decir que el determinismo físico fue una ensoñación de omnisciencia que parecía hacerse cada vez más real con cada avance de la física y que, finalmente, se convirtió en una pesadilla aparentemente ineludible. Sin embargo, las ensoñaciones de los psicólogos nunca pasaron de ser castillos en el aire: no eran más que sueños utópicos de alcanzar el mismo nivel que la física con sus métodos matemáticos y sus poderosas aplicaciones, incluso un nivel superior, al moldear hombres y sociedades. (Mientras que estos sueños totalitarios no son serios desde un punto de vista científico, son políticamente muy peligrosos [35]; ahora bien, como ya me he ocupado en otro lugar de esos peligros, no los discutiré aquí.)

IX

He dicho que el determinismo físico era una pesadilla. Lo es porque afirma que el mundo en su conjunto, con todo lo que hay en él, es un inmenso autómata y, por tanto, nosotros no somos más que diminutos engranajes o, a lo sumo, sub-autómatas suyos.

De este modo destruye, concretamente, la idea de creatividad. Reduce a una mera ilusión la idea de que al preparar esta conferencia haya utilizado mi cerebro para crear *algo nuevo*. Según el determinismo físico, lo úncio que ha ocurrido es que ciertas partes de mi cuerpo han hecho señales negras sobre un papel blanco: cualquier físico, con una información suficiente-

---

parar este pasaje con las págs. 404 y sigs. [págs. 192 y sigs.] (en las que Hume dice "Defino la necesidad de dos maneras,"), así como con su atribución a la "materia" de "esa cualidad inteligible, llamémosla o no necesidad" que, según dice, todo el mundo "ha de admitir que pertenece a la voluntad" (o "a las acciones de la mente"). En otras palabras, Hume intenta aquí aplicar a la "materia", es decir, a la física, su teoría del hábito o costumbre, así como su psicología asociacionista.

[35] Véase especialmente B. F. Skinner, *Walden Two*, 1948 [*Walden Dos*, Barcelona, Fontanella, 1968], un sueño utópico de omnipotencia realmente encantador y benevolente, aunque manifiestamente ingenuo (véase especialmente las págs. 246-50, así como la 214 y sigs. [traducción castellana, págs. 295-300 y 259 respectivamente]). Contra ello constituyen dos buenos antídotos. *Brave New World*, 1932, de Aldous Huxley [*Un mundo Feliz*, traducción castellana de Ramón Hernández, Barcelona, Plaza y Janés, 1963] (véase también *Brave New World Revisited*, 1959 [*Retorno al Mundo Feliz*, Sudamericana]), y *1984*, de George Orwell, 1948 [traducción castellana, *1948*, Barcelona, Destino, 1952; hay reedición en Libros de R.T.V., 1971]. He criticado algunas de estas ideas utópicas y autoritarias en *The Open Society and its Enemies*, 1945, 4.ª ed., 1962 [hay traducción castellana de Eduardo Loedel, *La Sociedad Abierta y sus Enemigos*. Buenos Aires, Paidos, 1957] y en *The Poverty of Historicism*, por ejemplo, la pág. 91 [véase la traducción ya citada, pág. 94 y sig.]. (Véase en ambos mi crítica a la llamada "sociología del conocimiento".)

mente detallada, podría haber escrito mi conferencia por el sencillo expediente de predecir los lugares exactos en que el sistema físico, constituido por mi cuerpo (que incluye evidentemente mi cerebro y mis dedos) y mi pluma, pondría esas marcas negras.

Pondré un ejemplo aún más impresionante: si el determinismo físico está en lo cierto, entonces un físico completamente sordo que nunca haya oído una composición musical podrá escribir todas las sinfonías y conciertos de Mozart o Beethoven por el simple expediente de estudiar los estados físicos exactos de sus cuerpos y predecir en qué lugar del pentagrama habrían de poner las señales negras. Nuestro físico sordo podría hacer aún más cosas: mediante un estudio suficientemente detallado de los cuerpos de Mozart o Beethoven podría componer partituras que ellos, de hecho, nunca escribieron, pero que habrían escrito si hubiesen sido distintas algunas circunstancias de sus vidas: si hubiesen comido cordero, pongo por caso, en lugar de pollo o si hubiesen bebido té en lugar de café.

Todo esto puede hacer nuestro físico sordo si se le suministra un conocimiento suficiente de las puras condiciones físicas. No necesitaría saber nada de la teoría musical —aunque debiera ser capaz de predecir todas las respuestas que hubiesen dado Mozart o Beethoven en un examen a base de preguntas relativas a la teoría del contrapunto.

Creo que todo esto es absurdo [36], y creo que el absurdo se hace aún más patente si aplicamos a un determinista este método de predicción física.

Según el determinismo, una teoría —como, por ejemplo, el determinismo— se sostiene a causa de cierta estructura física de su defensor (tal vez de su cerebro). De acuerdo con ello, nos estamos engañando a nosotros mismos (y estamos físicamente determinados a hacerlo) cuando creemos que existen cosas tales como argumentos o razones que nos hacen abrazar el determinismo. En otras palabras, si el determinismo físico es verdadero, no es defendible, ya que debe explicar todas nuestras reacciones (incluso las que nos parecen creencias basadas en argumentos) en términos de *condiciones puramente físicas*. Las condiciones puramente físicas, junto con nuestro medio físico, nos obligan a decir o aceptar lo que decimos o aceptamos. Un físico experimentado que no sepa una palabra de francés ni haya oído

---

[36] Mi físico sordo se asemeja mucho al demonio de Laplace, como es evidente (véase la nota 15). Creo que estas construcciones son absurdas por la sencilla razón de que los aspectos no físicos (fines, propósitos, tradiciones, gustos, ingenio) desempeñan algún papel en el desarrollo del mundo físico. En otras palabras, creo en el *interacionismo* (véanse las notas 44 y 64). Samuel Alexander, en su libro *Space, Time and Deity*, 1920, vol. II, pág. 328, dice de lo que él llama "calculador laplaciano": "La hipótesis del calculador es absurda, excepto en el sentido restringido apuntado". Sin embargo, el "sentido restringido" *incluye* la predicción de *todos* los acontecimientos puramente físicos, con lo cual habrá de *incluir* la posición de todas las marcas negras escritas por Mozart y Beethoven. Sólo *excluye* la predicción de experiencias mentales (exclusión muy semejante a mi suposición de la sordera del físico). Por tanto, Alexander está dispuesto a admitir eso que yo considero absurdo. (Tal vez deba decir aquí que prefiero discutir el problema de la libertad respecto a la creación musical o de nuevas teorías científicas que respecto a la ética o a la responsabilidad ética.)

hablar nunca del determinismo, habrá de poder predecir lo que diría un determinista francés en una discusión francesa acerca del determinismo, así como lo que diría su oponente indeterminista. Ahora bien, lo que esto significa es que si creemos haber aceptado una teoría como la determinista movidos por la fuerza lógica de ciertos argumentos, entonces según el determinismo físico nos estamos autoengañando; o, más exactamente, estamos en unas condiciones físicas que nos determinan a autoengañarnos.

Hume vio algo de esto, aunque al parecer no se percató del alcance que tenía para sus propios argumentos, ya que se limitó a comparar el determinismo de "*nuestros juicios*" con el de "*nuestras acciones*", diciendo que "*no tenemos mayor libertad en un caso que en el otro*" [37].

Este tipo de consideraciones tal vez expliquen por qué tantos filósofos no toman en serio el problema del determinismo físico, desestimándolo debido a su carácter "fantasmal" [38]. Sin embargo, en 1751, mucho antes de que la teoría de la evolución fuese generalmente aceptada, Lamettrie defendió del modo más vigoroso y serio la doctrina según la cual *el hombre es una máquina*. La teoría de la evolución confirió al problema una posición aún más fuerte, al sugerir que no hay una distinción clara entre la materia viva y la no viva [39]. A pesar de la victoria de la nueva teoría cuántica y de la conversión de tantos físicos al indeterminismo, la doctrina de Lamettrie, según la cual el hombre es una máquina, tal vez tenga hoy más partidarios que nunca entre los físicos, biólogos y filósofos; especialmente en la versión según la cual el hombre es un computador [40].

Si aceptamos una teoría evolucionista (como la de Darwin), entonces, aunque mostremos nuestro escepticismo acerca de la teoría de que la vida emerger a partir de la materia inorgánica, difícilmente podremos negar que ha tenido que haber una época en la que no existían —o, en todo caso, no tenían efecto alguno sobre el universo físico— las entidades abstractas de carácter no físico (tales como las razones, argumentos o conocimientos científicos) ni las reglas abstractas (tales como las reglas para la construc-

---

[37] Hume, *op cit., pág.* 609 (el subrayado es mío).

[38] Véase la nota 15 más arriba, así como el libro de Gilbert Ryle, *The Concept of Mind,* 1949, págs. 76 y sigs. [traducción castellana de Eduardo Rabossi, *El concepto de lo Mental,* Buenos Aires, Paidos, 1967; págs. 67 y sigs.] ("El fantasma del mecanicismo").

[39] Cf. N. W. Pirie, "The meaninglessness of the Term Life and Living", *Perspectives in Biochemistry,* 1937 (ed. J. Needham y D. E. Green), págs. 11 y sigs.

[40] Véase, por ejemplo, el artículo de A. M. Turing, "Computing Machinery and Intelligence", *Mind, 59,* 1960, págs. 433-60. Turing afirma que los hombres y las computadoras son indistinguibles en principio por sus funciones (comportamentales) observables y reta a sus oponentes a que *especifiquen* alguna realización o comportamiento observable del hombre que, en principio, no pueda llevar a cabo una máquina. Pero este desafío es una trampa intelectual: al *especificar* un tipo de comportamiento, suministramos condiciones para la construcción de un computador. Además, usamos y construimos computadores para que realicen aquellas cosas que nosotros no podemos hacer, del mismo modo que yo empleo una pluma o un lápiz para hacer una suma que me resulta imposible resolver de memoria. Einstein acostumbraba a decir "mi lápiz es más listo que yo", lo cual no demuestra que fuese indistinguible de su lápiz. (Cf. los párrafos finales de mi artículo sobre el indeterminismo citado más arriba en la nota 16, así como el capítulo 12, sección 5, de mi libro *Conjectures and Refutations.*)

ción de ferrocarriles, excavadoras, satélites artificiales o, también, las reglas gramaticales y las del contrapunto). Es difícil hacerse una idea de cómo podría producir el universo físico unas entidades abstractas, como las reglas, para caer luego bajo su control de tal modo que éstas tengan, a su vez, efectos palpables sobre el universo físico.

No obstante, hay una escapatoria a esta dificultad que, aunque tal vez resulte un tanto evasiva, en todo caso es fácil de tomar. Podemos negar sencillamente que existan dichas entidades abstractas, así como que puedan influir sobre el universo físico. Podemos decir también que lo que existen son nuestros cerebros que no son más que máquinas como las computadoras, que las supuestas reglas abstractas no son más que entidades físicas concretas mediante las que "programamos" las computadoras y que la existencia de cosas no físicas no es quizá más que "una ilusión" o, en todo caso, son cosas que carecen de importancia, puesto que todo marcharía igual aunque no existiesen tales ilusiones.

De acuerdo con esta escapatoria, no necesitamos preocuparnos por la condición "mental" de estas ilusiones. Puede que sean propiedades universales de las cosas: la piedra que he lanzado puede hacerse la ilusión de que salta, de la misma manera que yo me hago la ilusión de que la tiro. Del mismo modo, mi pluma o mi computadora puede tener la ilusión de que trabaja debido a su interés por los problemas que cree estar resolviendo —y que yo creo estar resolviendo— cuando en realidad lo único importante que está ocurriendo ahí es una serie de interacciones puramente físicas.

Como ustedes pueden ver con todo esto, el problema del determinismo que tanto preocupaba a Compton es realmente un problema serio. No es simplemente un embrollo filosófico, sino que también afecta al menos a los físicos, biólogos, conductistas, psicólogos, así como a los técnicos en computadoras.

Es evidente que unos pocos filósofos han tratado de mostrar (siguiendo a Hume o Schlick) que nuestro problema no es más que un embrollo lingüístico relativo al uso de la palabra "libertad". Sin embargo, estos filósofos no han logrado ver la diferencia existente entre el problema del determinismo físico y el filosófico: o bien son deterministas, como Hume, lo cual explica por qué para ellos la "libertad" no es "más que una palabra", o bien nunca han mantenido ese contacto íntimo con las ciencias físicas o con la técnica de las computadoras que les hubiese mostrado que tenían que vérselas con algo más que un simple embrollo lingüístico.

X

Como Compton, soy uno de los que toman en serio el problema del determinismo físico y, como él, no creo que seamos meras máquinas computadoras (aunque admito de buen grado que podemos aprender de ellas muchísimas cosas, incluso sobre nosotros mismos). Por tanto, soy, como Compton, un *indeterminista físico*: creo que el indeterminismo físico es un

requisito previo necesario para resolver nuestro problema. Hemos de ser indeterministas, aunque intentaré mostrar que el indeterminismo no basta.

Con este enunciado, *el indeterminismo no basta,* no me limito a señalar un nuevo aspecto, sino que he alcanzado el meollo mismo del problema que se puede explicar del modo siguiente:

Si el determinismo es verdadero, entonces el mundo en su conjunto es un reloj impecable que funciona con toda exactitud, lo mismo que las nubes, los organismos, los animales y los hombres. Por otro lado, si es verdadero el indeterminismo de Peirce o el de Heisenberg, o cualquier otro tipo de indeterminismo, entonces el puro *azar* desempeña un papel fundamental en nuestro mundo físico. *Ahora bien, ¿acaso el azar es realmente más satisfactorio que el determinismo?*

La cuestión es de sobra conocida. Los deterministas, como Schlick, la han expuesto de este modo: "...la libertad de acción, la responsabilidad y la salud mental no pueden trascender el dominio de la causalidad: se detienen donde comienza el azar ...un nivel superior de aleatoriedad... [sólo significa] un nivel superior de irresponsabilidad" [41].

Tal vez pueda explicar las ideas de Schlick mediante un ejemplo que ya he utilizado antes: decir que las marcas negras que he hecho sobre un papel blanco para preparar esta conferencia no eran sino el resultado del *azar,* no es más satisfactorio que afirmar su predeterminación física. De hecho, es aún menos satisfactorio. La razón de ello es que algunas personas estarán absolutamente dispuestas a creer que el texto de mi conferencia se puede explicar totalmente en términos de mi herencia y medios físicos, incluyendo mi educación, los libros que he leído y las charlas que he oído; pero es difícil que alguien crea que lo que les estoy leyendo sea sencillamente un resultado del azar; es decir, una muestra aleatoria de palabras o tal vez letras, unidas sin ningún propósito, deliberación, plan o intención.

La idea de que la única alternativa al determinismo es el puro azar (así como muchas de sus opiniones sobre el tema) la tomó Schlick de Hume quien decía que "la supresión" de lo que él llamaba "necesidad física" debe dar siempre como resultado "lo mismo que el *azar.* Como los objetos deben o no hallarse enlazados ...es imposible admitir un término medio entre el azar y la necesidad absoluta" [42].

Más adelante expondré argumentos en contra de esta importante doctrina según la cual el puro azar es la única alternativa al determinismo. Sin embargo, he de admitir que dicha doctrina parece aplicarse correctamente a los modelos de la teoría cuántica pensados para explicar, o al menos ilustrar, la posibilidad de la libertad humana. Parece ser ésta la razón por la cual estos modelos son tan insatisfactorios.

El propio Compton construyó un modelo de este tipo, aunque no le gustaba demasiado. Utilizaba la indeterminación cuántica y el carácter im-

---

[41] Véase M. Schlick, *Erkenntnis, 5,* pág. 183 (extractado de las ocho últimas líneas del primer párrafo).

[42] Hume, *op. cit.,* pág. 171 [trad. cit., pág. 275-6]. Véase, por ejemplo, la pág. 407 [ibid., pág. 189]: "...la libertad... es exactamente lo mismo que el azar".

predictible de los saltos cuánticos como modelo de decisión humana importante. Dicho modelo consiste en un amplificador que aumenta el efecto de un único salto cuántico de manera que o causa una explosión o destruye el relé necesario para producir la explosión. De este modo, un único salto cuántico puede equivaler a una decisión de mucho alcance. Pero en mi opinión, el modelo no se parece nada a una *decisión racional*, sino que se asemeja más bien al tipo de decisión que toman las personas cuando, al no poder deliberar, optan por decir: "echémoslo a cara o cruz". De hecho, todo este montaje para amplificar el salto cuántico parece un tanto innecesario: sirve lo mismo tirar al aire una moneda para decidir según el resultado, si apretamos o no el gatillo. Además, es evidente que disponemos de computadoras con dispositivos similares al lanzamiento de una moneda para producir resultados aleatorios cuando hagan falta.

Quizá se pueda decir que algunas de nuestras decisiones *son* como tirar una moneda: son decisiones precipitadas tomadas sin deliberar, puesto que a veces no disponemos de tiempo suficiente para ello. Los conductores o los pilotos tienen que tomar a veces decisiones precipitadas de este tipo y si son muy experimentados o tienen buena suerte, el resultado será satisfactorio; de lo contrario, no será así.

Admito que el modelo del salto cuántico pueda servir para esas decisiones precipitadas. Incluso admito la posibilidad de que de hecho, cuando tomamos una decisión precipitada, tenga lugar en nuestro cerebro algo similar a la amplificación de un salto cuántico. Ahora bien ¿son en realidad tan importantes las decisiones precipitadas? ¿Son características del comportamiento humano, del comportamiento humano *racional*?

No lo creo, ni pienso que vayamos a progresar mucho con los saltos cuánticos. Son precisamente esta clase de ejemplos los que parecen apoyar la tesis de Hume o Schlick de que el azar puro es la única alternativa al determinismo estricto. Lo que necesitamos para comprender el comportamiento humano racional —así como el animal— es algo que posea un carácter *intermedio* entre el azar perfecto y el determinismo perfecto —algo intermedio entre las nubes perfectas y los relojes perfectos.

La tesis ontológica de Hume y Schlick, según la cual no puede haber nada entre el azar y el determinismo, no sólo me parece dogmática (por no decir doctrinaria), sino también, claramente absurda; sólo se entiende si suponemos que creían en un determinismo completo en el que el azar no desempeñaba ninguna función, excepto la de ser síntoma de nuestra ignorancia. (Pero, a pesar de todo, me sigue pareciendo absurda, ya que es evidente que existe algo así como el conocimiento o ignorancia parcial.) Puesto que sabemos que incluso los relojes altamente fiables no son realmente perfectos, Schlick (si no Hume) debería de saber que ello se debe en gran medida a factores tales como la fricción —es decir, a efectos estadísticos o fortuitos. También sabemos que nuestras nubes no son perfectamente aleatorias, puesto que muchas veces podemos predecir el tiempo con éxito completo, al menos para períodos breves.

XI

Así pues, tendremos que retornar a nuestra vieja ordenación de nubes a la izquierda, relojes a la derecha y animales y hombres en algún punto intermedio.

Pero después de haberlo hecho (y hay que resolver algunos problemas antes de poder decir que esta ordenación es acorde con la física de nuestros días), aún entonces, lo único que hemos conseguido, en el mejor de los casos, es dar cabida a nuestra pregunta fundamental.

Como es obvio, lo que queremos es comprender de qué modo las cosas no físicas, como los *propósitos, deliberaciones, planes, decisiones, teorías, intenciones* y *valores* pueden tomar parte en la introducción de cambios físicos en el mundo, como es obvio que lo hacen, mal que les pese a Hume, Laplace y Schlick. Es evidentemente falso que todos esos tremendos cambios físicos que producen continuamente nuestras plumas, lápices o excavadoras se puedan explicar en términos exclusivamente físicos, sea mediante una teoría física determinista, sea mediante una teoría estocástica (según la cual se deberían al azar).

Compton era plenamente consciente de este problema, como lo demuestra este pasaje encantador sacado de sus conferencias Terry.

> Hace ya algún tiempo que escribí al secretario de la Universidad de Yale aceptando dar una conferencia el 10 de noviembre, a las cinco de la tarde. El secretario tenía tal confianza en mí que anunció públicamente mi asistencia; el auditorio, a su vez, confiaba tanto en su palabra que llegó a la sala a la hora señalada. Sin embargo, consideremos la gran improbabilidad física de que su confianza estuviese justificada. Entre tanto, mi trabajo me había llevado a las Montañas Rocosas y, más allá del océano, a la cálida Italia. [No es fácil] que un organismo fototrópico [como el mío] consintiese en abandonar aquel lugar para trasladarse al frío New-Haven. Las probabilidades de que yo estuviese en otra parte en aquel momento eran infinitas en número. Considerada desde el punto de vista de un acontecimiento físico, la probabilidad de cumplir mi compromiso era fantásticamente pequeña. ¿Por qué, pues, estaba justificada la creencia del auditorio?... Conocían mi intención y fue ella [la que] determinó que acudiese allí [43].

Compton muestra aquí con una gran belleza que el mero indeterminismo físico no basta. Está claro que hemos de ser indeterminista, pero hemos de intentar comprender también de qué modo los hombres y quizá los animales son susceptibles de verse "influidos" o "controlados" por cosas tales como fines, propósitos, reglas o acuerdos.

Este, pues, es nuestro problema central.

---

[43] Cf. *The Freedom of Man*, págs. 53 y sig.

XII

No obstante, un examen más detenido nos muestra que en la historia del viaje que hizo Compton de Italia a Yale hay *dos* problemas. Al primero lo llamaré *problema de Compton* y al segundo, *problema de Descartes*.

Pocas veces se han percatado los filósofos del problema de Compton y, cuando ha sido así, lo han visto de manera poco clara. Se puede formular como sigue:

Existen cosas tales como cartas en las que se aceptan propuestas para dar conferencias y existen anuncios públicos de las intenciones; hay declaraciones públicas de fines y propósitos; hay reglas morales generales. Todos estos documentos, declaraciones o reglas disponen de un contenido o significado que permanece invariable a través de traducciones o reformulaciones. Por tanto, *dicho contenido o significado es algo totalmente abstracto*. Con todo, puede ejercer un control —tal vez mediante una anotación críptica en una agenda de compromisos— sobre los movimientos físicos de las personas hasta el punto de arrancarlas de Italia y enviarlas a Connecticut. ¿Cómo es posible esto?

Es a esto a lo que llamo problema de Compton. Es importante notar que, planteado así, el problema es indiferente respecto a la adopción de una psicología mentalista o conductista: en la formulación aquí presentada y sugerida por el texto de Compton, el problema se formula en términos de su *conducta* consistente en volver a Yale; pero habría muy poca diferencia si hubiésemos introducido acontecimientos mentales, como voliciones o el sentimiento de haber captado o cogido una idea.

Manteniendo su propia terminología conductista, el problema de Compton puede describirse como el problema de la influencia del *universo de los significados abstractos* sobre la conducta humana (y por ende, sobre el universo físico). Aquí "universo de los significados" es una abreviatura de cosas muy diversas, como promesas, fines y diversos tipos de reglas como las gramaticales, las de urbanidad, las de la lógica, las del ajedrez o las del contrapunto; sin olvidar cosas tales como publicaciones científicas (y de otro tipo), llamamientos a nuestro sentido de justicia, de generosidad o a nuestro gusto artístico, etc., etc., casi *ad infinitum*.

Creo que lo que llamo aquí problema de Compton, aunque pocos filósofos lo hayan visto, es uno de los más interesantes de la filosofía. En mi opinión, es realmente un problema clave mucho más importante que el clásico problema del cuerpo y la mente que denomino aquí "problema de Descartes".

Para evitar malentendidos, tal vez deba mencionar que Compton no pretendía suscribir un conductismo plenamente desarrollado cuando formuló su problema en aquellos términos. Por el contrario, no dudaba de la existencia de su propia mente ni de la de otras mentes, ni de experiencias como las voliciones, deliberaciones, el placer o el dolor. Por tanto, habría puesto de relieve un *segundo* problema a resolver.

Podemos identificar este segundo problema con el problema clásico del cuerpo y la mente o problema de Descartes. Se puede formular así: ¿Cómo es posible que los estados mentales —voliciones sentimientos, expectativas— influyan o controlen los movimientos físicos de nuestras piernas? O, también (aunque en este contexto es menos importante), ¿cómo es posible que los estados físicos de un organismo puedan influir sobre sus estados mentales? [44].

Compton sugiere que cualquier solución *satisfactoria aceptable* a alguno de los dos problemas ha de satisfacer el siguiente postulado que denominaré *postulado de libertad de Compton*: la solución ha de explicar la libertad, así como de qué manera la libertad no se reduce al azar, sino que es más bien el resultado de una sutil interrelación entre *algo casi aleatorio o fortuito* y *algo así como un control selectivo o restrictivo* —como puede ser un fin o una norma—, aunque no se trate de un control férreo. Es evidente que los controles que guiaban la vuelta de Compton de Italia le permitían una plena libertad: libertad, digamos, de elegir entre un buque americano, francés o italiano o la de posponer su conferencia si hubiese surgido una obligación más importante.

Podemos decir que el postulado de libertad de Compton reduce las soluciones aceptables de nuestros dos problemas, exigiendo que se acomoden a *la idea de combinar libertad y control*, así como a *la idea de "control plástico"*, frente a "control férreo".

El postulado de Compton constituye una restricción que acepto gustosa y libremente. Además, mi aceptación crítica, libre y deliberada de esta restricción puede tomarse como ejemplo de esa combinación de libertad y control que constituye el contenido mismo del postulado de libertad de Compton.

XIII

He explicado nuestros dos *problemas* centrales —el de Compton y el de Descartes—. Para resolverlos creo que precisamos una *nueva teoría*. Se trata, de hecho, de una nueva teoría evolucionista y de un nuevo modelo de organismo.

Esta necesidad surge porque las teorías indeterministas existentes son

---

[44] En los capítulos 12 y 13 de mi libro *Conjectures and Refutations* se encontrará una discusión crítica de lo que denomino aquí problema de Descartes. He de decir aquí que, como Compton, soy casi cartesiano en la medida en que rechazo la tesis del carácter físicamente completo de los organismos vivos (considerados como sistemas físicos); es decir, en la medida en que conjeturo que, en algunos organismos, los estados mentales pueden entrar en *interacción* con los estados físicos. (No obstante, soy menos cartesiano que Compton: me siento aún menos atraído que él por los modelos en términos de conmutadores generales; cf. notas 45, 46 y 64.) Además, no siento ninguna simpatía por lo que dice Descartes acerca de una *substancia* mental o pensante —ni tampoco por su *substancia* material o extensa. Soy cartesiano sólo en la medida en que creo en la existencia tanto de *estados* físicos como de *estados* mentales (y, además, en cosas mucho más abstractas, como el estado de una discusión).

insatisfactorias. Son indeterministas, pero sabemos que el indeterminismo no basta, pues no está claro de qué modo elude la objeción de Schlick o de qué manera cumple el postulado de *libertad más control* de Compton. Una vez más, el problema de Compton las trasciende: no tienen mayor relevancia para dicho problema. Aunque dichas teorías constituyan intentos de resolver el problema de Descartes, las soluciones que ofrecen no parecen satisfactorias.

Las teorías a que aludo pueden denominarse "modelos de control de conmutador general" o, brevemente, "teorías de conmutador general". La idea básica es que nuestro cuerpo es una especie de máquina que puede regularse mediante una llave o conmutador desde uno o más *puntos centrales de control.* Descartes llegó incluso a situar exactamente el punto de control: decía que la acción de la mente sobre el cuerpo se ejercía en la glándula pineal. Algunos teóricos cuánticos han sugerido (sugerencia que Compton aceptó de modo muy provisional) que nuestras mentes actúan sobre el cuerpo influyendo sobre —o seleccionando— algunos saltos cuánticos. Estos se ven entonces amplificados por nuestro sistema nervioso central que funciona como un amplificador electrónico: los saltos cuánticos amplificados funcionan como una cascada de relés o conmutadores generales que terminan desencadenando contracciones musculares [45]. Creo que en los libros de Compton hay síntomas que indican que no le agradaba mucho esta teoría o modelo concreto que utilizaba con un único fin: mostrar que el indeterminismo humano (o incluso, la "libertad") no contradice necesariamente la física cuántica [46]. Creo que estaba en lo cierto en este punto, así como en su insatisfacción frente a las teorías del conmutador general.

La razón estriba en que estas teorías —sea la cartesiana o las del amplificador de los físicos cuánticos— pertenece a lo que tal vez podríamos denominar "*teorías del niño pequeñito*". Me parecen casi tan poco atractivas como los *niños pequeñitos.*

Estoy seguro de que todos ustedes conocen la historia de la madre soltera que alegaba: "Pero si sólo es un niño *muy* pequeñito". Las protestas de Descartes me parecen similares: "Pero si es tan pequeño: no es más que un punto matemático inextenso aquel en que la mente actúa sobre el cuerpo".

Los teóricos cuánticos mantienen una teoría del niño pequeño muy similar: "Pero si es sólo mediante *un* salto cuántico y precisamente con las incertidumbres de Heisenberg —que también son muy pequeñas— con lo que la mente actúa sobre un sistema físico." Admito que aquí puede haber

---

[45] Compton discutió con algún detalle esta teoría, especialmente en *The Freedom of Man,* págs. 37-65. Véase especialmente la referencia a Ralph Lillie, *op. cit.,* en *The Freedom of Man,* pág. 50. Véase también *The Human Meaning of Science,* pág. 53, acerca *del carácter de individualidad de nuestras acciones* y la explicación de por qué nos permite eludir lo que podríamos llamar el segundo cuerno del dilema (el primero es el determinismo puro); es decir, la posibilidad de que nuestras acciones se deban al *puro azar;* cf. la nota 41.

[46] Véase especialmente *The Human Meaning of Science,* págs. VIII y sigs., y página 54, el último enunciado de la sección.

un ligero progreso en la medida en que se especifica el tamaño del bebé. Pero sigue sin gustarme la criatura.

Por pequeño que sea el conmutador general, el modelo del tal conmutador *cum* amplificador sugiere con gran fuerza que todas nuestras decisiones o son precipitadas (como ya he señalado en la sección X) o se componen de decisiones precipitadas. Ahora bien, admito que los mecanismos amplificadores constituyen una característica importante de los sistemas biológicos, ya que la energía de reacción liberada o desencadenada por un estímulo biológico, normalmente excede con mucho la energía del estímulo desencadenador [47]. También admito, evidentemente, que hay decisiones precipitadas. Pero difieren claramente del tipo de decisión en que Compton pensaba: son casi como los reflejos, con lo cual no se adecúan ni a la situación del problema de Compton relativa a la influencia del universo del significado sobre nuestro comportamiento, ni a su postulado de libertad (ni a la idea de control "plástico"). Por regla general, las decisiones que se adecúan a todas estas cosas se toman de un modo casi imperceptible mediante una *deliberación* prolongada. Se toman mediante una especie de proceso de *maduración* que no está bien recogido en el modelo del conmutador general.

Si tomamos en cuenta este proceso de deliberación, podemos sacar otra sugerencia para nuestra nueva teoría, ya que la deliberación funciona siempre por *ensayo y error* o, más exactamente, por *el método de ensayo y supresión de errores*: proponiendo diversas posibilidades de un modo provisional y eliminando aquellas que no parecen adecuadas. Esta situación sugiere que en nuestra nueva teoría hemos de emplear algún mecanismo de ensayo y supresión de errores.

Voy a bosquejar el procedimiento que pretendo seguir antes de formular en términos generales mi teoría evolucionista.

Mostraré primero cómo funciona en un caso particular, aplicándola a nuestro primer problema; es decir, al problema de Compton relativo a la *influencia del significado sobre el comportamiento.*

Tras haber resuelto el problema de Compton de esta manera, formularé la teoría en términos generales. A continuación, veremos que también contiene —en el marco de nuestra teoría, que crea una nueva situación proble-

---

[47] Es este un punto de gran importancia; tanto, que difícilmente podremos decir que un proceso es típicamente biológico si no incorpora la liberación o descarga de energía almacenada. Pero, evidentemente, no se puede decir lo contrario: muchos procesos no biológicos tienen el mismo carácter. Aunque los procesos amplificadores o desencadenadores no desenpeñen un gran papel en la física clásica, son muy característicos de la física cuántica, así como de la química. (La radioactividad, con una energía de descarga igual a cero, es un caso extremo; otro caso interesante —en principio adiabático— es la sintonización con una determinada frecuencia de radio, seguida de la tremenda amplificación de la señal o estímulo.) Por esta razón entre otras, fórmulas tales como "la causa es igual al efecto" (y, con ella, la crítica tradicional del interaccionismo cartesiano) han pasado de moda hace ya tiempo, a pesar de la validez permanente de las leyes de conservación. Cf. la nota 44 y la función *estimuladora o desencadenadora* del lenguaje, discutida en la sección XIV, más adelante; véase también mi libro *Conjectures and Refutations*, pág. 381 [trad. cit., págs. 438-9].

mática— una solución directa y casi trivial al problema cartesiano clásico relativo al cuerpo y la mente.

XIV

Abordemos ahora el primer problema —es decir, el problema de Compton relativo a la influencia del significado sobre la conducta— sirviéndonos de algunos comentarios en torno a *la evolución de los lenguajes, desde los lenguajes animales hasta los humanos.*

Los lenguajes animales tienen con los humanos muchas cosas en común; pero también existen diferencias: como todos sabemos, éstos últimos trascienden de algún modo los lenguajes animales.

Utilizando y ampliando algunas de las ideas de mi último maestro Karl Bühler [48], distinguiré dos funciones compartidas por los lenguajes animales y humanos y otras dos que sólo poseen los humanos; en otras palabras, distinguiré dos funciones inferiores y dos superiores que han evolucionado sobre la base de las inferiores.

Las dos funciones inferiores del lenguaje son las siguientes. En primer lugar, el lenguaje como toda otra forma de comportamiento, consta de *síntomas o expresiones;* es sintomático o expresivo del estado del organismo que hace los signos lingüísticos. Siguiendo a Bühler, llamo a esto *la función sintomática o expresiva del lenguaje.*

En segundo lugar, para que tenga lugar el lenguaje o la comunicación, no sólo debe haber un organismo que haga señales (el "transmisor"), sino que también tiene que haber otro que reaccione (el "receptor"). La expresión sintomática del primer organismo, el transmisor, desencadena, evoca, estimula o dispara una reacción en el segundo que *responde* a la conducta del transmisor, convirtiéndola así en una *señal.* Bühler llamaba a esta función lingüística consistente en actuar sobre un receptor, *la función desencadenadora o señalizadora del lenguaje.*

Para poner un ejemplo, un pájaro que esté presto a salir volando puede *expresar* su disposición exhibiendo ciertos síntomas que pueden *desencadenar o disparar* ciertas respuestas o reacciones en un segundo pájaro, como consecuencia de lo cual puede a su vez estar presto a salir volando.

Nótese que ambas funciones, la expresiva y la desencadenadora, son

---

[48] La teoría de las funciones del lenguaje se debe a Karl Bühler (*The Mental Development of Child,* 1919, traducción inglesa, 1930; págs. 55, 56, 57 [traducción castellana de Rosario Fuentes, *El Desarrollo Espiritual del Niño,* Madrid, Espasa-Calpe, 1934]; véase también la *Sprachtheorie,* 1934 [traducción castellana de Julián Marías, *Teoría del Lenguaje,* Madrid, Revista de Occidente, 1967]). A sus tres funciones he añadido la argumentadora (así como algunas otras que no desempeñan aquí ninguna función, como son las funciones exhortativas y persuasivas). Véase, por ejemplo, mi artículo "El lenguaje y el problema del cuerpo y la mente", en *Conjectures and Refutations,* pág. 295, nota y texto [trad. cit., pág. 341]. (Véanse también las págs. 134 y siguientes [pág. 158].) No se puede decartar que en los animales, especialmente las abejas, haya estados de transición hacia lenguajes descriptivos; véase K. von Frisch, *Bees: their Vision, Chemical Senses and Language,* 1950; *The Dancing Bees,* 1955, así como M. Lindauer, *Communication Among Social Bees,* 1961.

distintas, ya que puede tener lugar la primera sin la segunda, aunque no a la inversa: mediante su comportamiento, un pájaro puede expresar su disposición al vuelo sin influir con ello en otro pájaro. Por tanto, la primera función puede desarrollarse sin la segunda, lo cual pone de manifiesto que ambas pueden separarse a pesar de que en los casos auténticos de comunicación lingüística siempre se den juntas.

Ambas funciones inferiores, la sintomática o expresiva, por un lado, y la desencadenadora o señalizadora, por el otro, son comunes a los lenguajes animales y humanos, estando siempre presentes cuando tienen lugar las funciones superiores (típicamente humanas).

El lenguaje humano es mucho más rico. Posee diversas funciones y dimensiones que los lenguajes animales no tienen. Dos de dichas funciones nuevas son de la mayor importancia para la evolución del razonamiento y la racionalidad: se trata de las *funciones descriptiva y argumentadora.*

Puedo ponerles a ustedes como ejemplo de función descriptiva la descripción de cómo hace dos días en mi jardín florecía una magnolia y lo que ocurrió luego cuando empezó a caer la nieve. De este modo, podría expresar mis sentimientos, desencadenando o disparando en ustedes otros sentimientos: puede que ustedes reaccionen pensando en *sus* magnolias. De este modo, ambas funciones inferiores están presentes. Pero, *además* de esto, les habré descrito algunos hechos, les habré formulado algunos *enunciados descriptivos* que de hecho pueden ser *verdaderos* o *falsos.*

Cuando hablo, no puedo dejar de expresarme y si ustedes me oyen, difícilmente podrán dejar de reaccionar. Por tanto, las funciones inferiores están *siempre* presentes. Con todo, *no es necesario* que la función descriptiva esté presente, ya que puedo hablarles sin describir ningún hecho. Por ejemplo, no necesito describir nada para mostrar o expresar incomodidad —la duda, pongo por caso— acerca de si ustedes lograrán sobrevivir a esta conferencia interminable. Con todo, la descripción, incluso la descripción del estado de cosas conjeturado que formulamos en forma de teorías o hipótesis, es sin duda una de las funciones más importantes del lenguaje humano. Es, además, la que distingue de la manera más tajante el lenguaje humano de los lenguajes animales (si bien en el lenguaje de las abejas parece haber algo muy próximo a la descripción [49]). Naturalmente, se trata de una función indispensable para la ciencia

La última y más elevada de las funciones que mencionaré en esta visión general es la *función argumentadora del lenguaje* que puede observarse, en su forma superior de desarrollo, en una *discusión crítica* disciplinada.

La función argumentadora del lenguaje, no sólo es la más elevada de las cuatro que discuto, sino que además es la última que aparece en el proceso evolutivo. Su evolución se ha conectado estrechamente con la de la actitud argumentadora, crítica y racional; y puesto que dicha actitud ha conducido a la evolución de la ciencia, podemos decir que la función argumentadora del lenguaje ha creado el instrumento tal vez más poderoso

[49] Cf. los libros de Frisch, *op. cit.*, y Lindauer, *op. cit.*

de adaptación biológica que haya surgido nunca en el transcurso de la evolución orgánica.

Al igual que las otras funciones, el arte de la argumentación crítica se ha desarrollado mediante el método de ensayo y supresión de errores y ha tenido una influencia de lo más decisiva sobre la capacidad humana de pensar racionalmente. (La propia lógica formal se puede considerar como un "organon de argumentación crítica") [50]. Al igual que el descriptivo, el uso argumentador del lenguaje ha hecho aparecer normas ideales de control o "*ideas reguladoras*" (para utilizar la terminología kantiana): la idea reguladora fundamental del uso descriptivo del lenguaje es *la verdad* (frente a la *falsedad*) y la del uso argumentador en la discusión crítica es *la validez* (frente a *la invalidez*).

Por regla general, los argumentos se orientan en pro o en contra de una proposición o enunciado descriptivo, lo cual explica por qué nuestra cuarta función —la argumentadora— tiene que haber surgido más tarde que la descriptiva. Cuando aduzco en una junta que la Universidad no debe aprobar determinado presupuesto porque carecemos de fondos o porque es más beneficioso gastar el dinero en otra cosa, no argumento solamente en pro o en contra de una *propuesta,* sino además, en pro o en contra de una *proposición;* en pro, digamos, de que el presupuesto proyectado no va a ser beneficioso y *en contra* de la proposición que afirma que dicho presupuesto dará buenos resultados. Así pues, los argumentos, incluso los argumentos relativos a propuestas, versan, por regla general, acerca de proposiciones y, muy frecuentemente, acerca de proposiciones *descriptivas.*

Sin embargo, el uso argumentador del lenguaje se puede separar tajantemente del uso descriptivo, por la sencilla razón de que es posible describir sin argumentar; es decir, puedo describir sin dar razones en pro o en contra de la verdad de mi descripción.

Nuestro análisis de las cuatro funciones del lenguaje —la expresiva, la señalizadora, la descriptiva y la argumentadora— puede resumirse diciendo que, aunque haya que admitir que las dos funciones inferiores —la expresiva y la señalizadora— están *siempre* presentes cuando tienen lugar las superiores, no obstante, tenemos que distinguir unas de otras.

Sin embargo, muchos conductistas y filósofos han pasado por alto las funciones superiores, al parecer a causa de que las inferiores están siempre presentes independientemente de las superiores.

XV

Al margen de las nuevas funciones del lenguaje que han evolucionado y surgido juntamente con el hombre y con la racionalidad humana, hemos de

[50] Véase mi libro *Conjectures and Refutations,* capítulo I, especialmente las consideraciones que hago en la pág. 64 [trad. cit., pág. 79] acerca de la lógica formal como "el *organon de la crítica racional*"; véanse también los capítulos del 8 al 11, así como el 15.

tener en cuenta otra distinción de una importancia similar: la distinción entre la evolución de los *órganos* y la de los *instrumentos o máquinas*, distinción que hay que atribuir a uno de los mayores filósofos ingleses, Samuel Butler, el autor de *Erewhon* (1872).

La Evolución de los animales progresa en gran medida, aunque no exclusivamente, mediante la modificación de los órganos (o la conducta) o mediante la emergencia de órganos nuevos *externos a nuestros cuerpos o personas*: "exosomáticamente" o "extrapersonalmente", como dicen los biólogos. Estos nuevos órganos son herramientas, armas, máquinas o casas.

Naturalmente, podemos encontrar en los animales inicios rudimentarios de dicho desarrollo exosomático. La construcción de cubiles, cuevas o nidos constituyen logros incipientes. Recuerden ustedes que los castores construyen diques muy ingeniosos. Pero el hombre, en lugar de desarrollar mejores ojos y oídos, produce gafas, microscopios, telescopios, teléfonos y aparatos para oír mejor; en lugar de desarrollar piernas más veloces, produce coches cada vez más rápidos.

Pero es otro el tipo de evolución exosomática o extrapersonal que ahora me interesa subrayar: en lugar de desarrollar una memoria mejor o un cerebro mejor, producimos papel, lápices, plumas, máquinas de escribir, dictáfonos, imprentas y bibliotecas.

Podemos considerar que estas adiciones a nuestro lenguaje —sobre todo a sus funciones argumentadora y descriptiva— constituyen dimensiones nuevas. El último eslabón de este desarrollo (utilizado fundamentalmente para apoyar nuestra capacidad argumentadora) es la aparición de las computadoras.

XVI

¿Cómo se relacionan las funciones y dimensiones superiores con las inferiores? No sustituyen las inferiores, como hemos visto, sino que ejercen sobre ellas una especie de *control plástico* —un control con retroalimentación.

Tomemos a guisa de ejemplo la discusión de una reunión científica. Puede ser estimulante y deleitable, provocando síntomas de que efectivamente es así. A su vez, dichas expresiones pueden desencadenar síntomas semejantes en otros participantes. Con todo, no cabe duda de que, hasta cierto punto, estos síntomas y señales desencadenadoras se deberán al *contenido* científico de la discusión que ejercerá un control sobre ellos. Puesto que dicho contenido será de *naturaleza descriptiva y argumentadora*, las funciones inferiores estarán controladas por las superiores. Además, aunque un buen chiste o un gesto divertido puedan hacer a corto plazo que ganen las funciones inferiores, a la larga lo que cuenta es un buen argumento —un argumento válido— con todo lo que demuestra o refuta. En otras palabras, nuestra discusión se ve controlada, aunque de un modo plástico, por las ideas reguladoras de verdad y validez. El descubrimiento y desarrollo de las nuevas dimensiones de impresión y publicación refuerza esta situación,

sobre todo cuando dichas dimensiones se emplean para imprimir y publicar tanto teorías e hipótesis científicas, como artículos en que se discuten críticamente.

No me es posible hacer aquí justicia a la importancia de los argumentos críticos: es un tema sobre el que he escrito muchas cosas [51] que no voy a repetir ahora. Sólo deseo hacer hincapié en el hecho de que los argumentos críticos son *medios de control*: son medios para eliminar errores, medios de selección. *Resolvemos nuestros problemas* proponiendo tentativamente diversas teorías e hipótesis rivales a modo de globos sonda, como si dijésemos, que sometemos a discusión crítica a fin de suprimir errores.

La evolución de las funciones superiores del lenguaje, que he tratado de describir, puede definirse como la evolución de nuevos medios de resolver problemas mediante nuevos tipos de ensayos y nuevos métodos de suprimir errores; es decir, nuevos métodos de *controlar* los ensayos.

XVII

Puedo ahora presentar mi solución a nuestro problema fundamental, es decir, al problema de Compton acerca de la influencia del significado sobre la conducta. Hela aquí.

Los niveles superiores del lenguaje han evolucionado bajo la presión de la necesidad de un *control más eficaz* de las cosas: nuestros niveles lingüísticos inferiores y nuestra adaptación al medio, utilizando para ello el método de desarrollar, no sólo herramientas nuevas, sino también, por ejemplo, nuevas teorías científicas y nuevas normas de selección.

Ahora bien, al desarrollar sus funciones superiores, nuestro lenguaje también ha desarrollado contenidos y significados abstractos; es decir, hemos aprendido a abstraer diversos métodos de formular o expresar teorías, así como a prestar atención a su *contenido o significado invariante* (del que depende su verdad). Todo esto no sólo se aplica a las teorías y otros enunciados descriptivos, sino también a propósitos y fines o a cualquier cosa susceptible de ser sometida a discusión crítica.

Denomino "problema de Compton" al problema de explicar y comprender el poder de control que poseen los significados, como son los contenidos de nuestras teorías, propósitos o fines; propósitos o fines que en algunos casos podemos adoptar tras una deliberación y discusión. Pero esto ya no es un problema. Su poder de influir sobre nosotros forma parte de esos contenidos y significados, ya que la función de los contenidos y significados es, en parte, controlar.

Esta solución del problema de Compton salva su postulado restrictivo, ya que el control que nuestras teorías y propósitos ejercen sobre nosotros

---

[51] Véase la nota 50 y mi libro *The Open Society and Its Enemies*, especialmente el capítulo 24 y el *Apéndice* al volumen II (4.ª edición, 1962), así como *Conjectures and Refutations*, sobre todo el Prefacio y la Introducción.

y nuestros actos es un control *plástico*. No nos vemos *forzados* a someternos al control de nuestras teorías, ya que podemos discutirlas críticamente y rechazarlas libremente si pensamos que no satisfacen nuestras normas reguladoras. Por tanto, el control no es, ni mucho menos, unidireccional. No sólo nos controlan nuestras teorías, sino que además nosotros podemos controlarlas a ellas (e incluso a nuestras normas): hay aquí una especie de retroalimentación. Si nos sometemos a nuestras teorías, lo hacemos libremente tras una deliberación; es decir, tras haber discutido críticamente las alternativas y tras haber elegido libremente entre las teorías rivales a la luz de la discusión crítica.

Presento esto como solución al problema de Compton; y antes de proceder a resolver el problema de Descartes, trataré brevemente la teoría general de la evolución que he utilizado ya implícitamente en mi solución.

XVIII

Presento mi teoría general con muchas disculpas. Me ha llevado mucho tiempo formularla plenamente y verla con claridad. Sin embargo, no estoy en absoluto satisfecho de ella. Esto se debe, en parte, al hecho de que se trata de una teoría *evolucionista* que temo que no añada gran cosa a las teorías evolucionistas existentes, exceptuando quizás un nuevo acento.

Me avergüenza confesarlo, porque cuando era más joven acostumbraba a hablar muy despreciativamente de las filosofías evolucionistas. Cuando hace veintidós años Canon Charles E. Raven dijo en su libro *Science, Religion and the Future* que la controversia darwinista había sido "una tormenta en una taza de té victoriana" me mostré de acuerdo, si bien lo critiqué [52] por prestar demasiada atención "a los vapores que aún salen de la taza", con lo que me refería al clima agobiante de las filosofías evolucionistas (especialmente aquellas que afirman que hay leyes inexorables de evolución). Mas he de confesar que esa taza de té se ha convertido, después de todo, en *mi* taza de té y con ella he de comerme mi humilde pastel.

Totalmente al margen de las *filosofías* evolucionistas, lo chocante de las *teorías* evolucionistas es su carácter tautológico o causi-tautológico: la dificultad estriba en que el darwinismo y la selección natural, a pesar de su inmensa importancia, explican la evolución mediante "la supervivencia del más apto" (expresión debida a Herbert Spencer). Sin embargo, no parece haber mucha diferencia —si es que la hay— entre decir "los que sobreviven son los más aptos" y la tautología "los que sobreviven son los que sobreviven". Esto es así porque me temo que no haya más criterio de aptitud que la supervivencia efectiva, de manera que del hecho de que haya sobrevivido un organismo concluimos que era el más apto o el más adaptado a las condiciones vitales.

---

[52] Cf. la pág. 106, nota 1, de mi libro *The Poverty of Historicism* [trad. cit., página 132, nota 2].

Esta situación muestra que el darwinismo, con todas sus grandes virtudes, dista de ser una teoría perfecta. Precisa una reformulación que la haga más precisa. La teoría evolucionista que voy a pergeñar constituye un intento de reformularla en dicho sentido.

Se puede definir mi teoría como un intento de aplicar a la evolución en su conjunto lo que hemos sacado en limpio del análisis de la evolución, desde los lenguajes animales hasta el lenguaje humano. Se trata de una *visión de la evolución* como un sistema de controles plásticos en desarrollo, así como de una *visión de los organismos* como elementos que incorporan —o, en el caso del hombre, que desarrollan exosomáticamente— ese sistema jerárquico de controles plásticos. Suponemos aquí la teoría neodarwinista de la evolución, aunque se reformula señalando que sus "mutaciones" pueden interpretarse como gambitos de ensayo y error más o menos accidentales y la "selección natural", como un modo de controlarlos mediante la supresión de errores.

Formularé ahora la teoría en forma de doce tesis breves.

(1) Todos los *organismos acometen* constantemente, día y noche, *la resolución de problemas,* cosa que también hacen todas las *secuencias* evolucionistas *de organismos* —los *phyla* que empiezan por las formas más primitivas y cuyos últimos miembros son los organismos vivos en la actualidad.

(2) Estos problemas son problemas en sentido objetivo: hipotéticamente, pueden ser reconstruidos retrospectivamente, como si dijéramos. (Me referiré de nuevo a esto, más adelante, con mayor detalle.) Los problemas objetivos, en este sentido, no precisan una contrapartida consciente y cuando la tienen, el problema consciente no tiene por qué coincidir con el problema objetivo.

(3) La resolución de problemas procede siempre por ensayo y error: se lanzan provisionalmente nuevas formas, nuevos órganos, nuevos modos de comportamiento y nuevas hipótesis que se controlan mediante eliminación de errores.

(4) La eliminación de errores puede proceder o bien por eliminación completa de las formas que no tienen éxito (la selección natural elimina las formas sin éxito) o bien por evolución (tentativa) de controles que modifiquen o supriman los órganos, formas, conductas o hipótesis sin éxito.

(5) El organismo individual incorpora [53] en sí mismo, como si dijéra-

---

[53] La idea de "incorporación" ["*telescoping*"] (aunque no el término, que debo a Alan Musgrave) tal vez pueda encontrarse en el capítulo VI del *Origin of Species,* 1859, de Charles Darwin (cito la edición Mentor Book., pág. 180; el subrayado es mío [cf. *El Origen de las Especies,* Barcelona, Zeus, 1970]): "...todo organismo de un alto nivel de desarrollo ha sufrido muchos cambios; y... toda estructura modificada tiende a heredarse, de modo que cada una de las modificaciones no ...se pierde totalmente... *Por tanto, la estructura de cada una de las partes* [del organismo] *...es la suma* de muchos cambios heredados por los que ha pasado la especie..." Véase también E. Baldwin en el libro *Pespectives in Biochemistry,* págs. 99 y sigs., así como la bibliografía que cita.

mos, los controles desarrollados a lo largo de la evolución de su *phylum* —como si reiterase parcialmente la evolución filogenética en su desarrollo ontogenético.

(6) El organismo individual es una especie de punta de flecha de la secuencia evolucionista de organismos a que pertenece (su *phylum*): él mismo es una solución tentativa que prueba nuevos nichos ecológicos, eligiendo y modificando el medio. Mantiene con su *phylum* unas relaciones casi exactas a las que las acciones (comportamiento) del organismo individual mantienen con éste: tanto el organismo individual como su comportamiento son ensayos que se pueden eliminar mediante la supresión de errores.

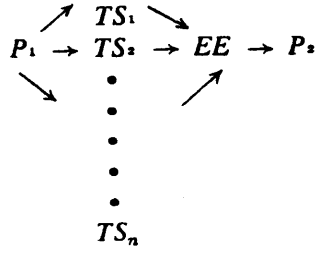
(7) Siendo "*P*" un problema, "*TS*" las soluciones tentativas y "*EE*" la eliminación de errores, podemos expresar del modo siguiente la secuencia evolutiva fundamental de los acontecimientos:

$$P \rightarrow TS \rightarrow EE \rightarrow P.$$

Ahora bien, esta secuencia no es circular: por lo general, el segundo problema es distinto del primero; es el resultado de una nueva situación que ha surgido, en parte, gracias a las soluciones tentativas que se han ensayado y a la eliminación de errores que las controlan. Para señalarlo, hemos de escribir de nuevo el esquema del siguiente modo:

$$P_1 \rightarrow TS \rightarrow EE \rightarrow P_2.$$

(8) Pero aún así, falta un elemento importante: la multiplicidad de las soluciones tentativas, la multiplicidad de los ensayos. Por tanto nuestro esquema final queda más o menos así:



Conocimiento básico

(9) De este modo, podemos comparar nuestro esquema con el del neodarwinismo. Según éste hay *un* problema fundamental: el de la supervivencia. Como en nuestro esquema, hay una multiplicidad de soluciones tentativas —las variaciones o las mutaciones—, aunque sólo hay *un* medio de eliminar errores —la muerte del organismo—. Además (en parte por esta razón) el hecho de que P_1 y P_2 difieran esencialmente o se pasa por alto o no se reconoce de un modo suficientemente claro su importancia fundamental.

(10) En nuestro sistema, no todos los problemas son de supervivencia: hay problemas y subproblemas muy específicos (aún cuando los problemas primitivos hayan sido simplemente problemas de supervivencia). Un problema primitivo, P_1, puede ser, por ejemplo, la reproducción. Su solución puede llevar a un nuevo problema P_2: el problema de dispersar o librarse de la prole —los hijos que amenazan con acabar no sólo con el organismo paterno, sino también con los suyos propios [54].

Tal vez sea interesante constatar que *el problema de evitar verse eliminado por la propia prole* puede ser uno de los problemas resueltos por la evolución de los *organismos multicelulares*: en lugar de librarse de la prole, se establece una economía común con diversos métodos nuevos de vivir juntos.

(11) La teoría aquí propuesta distingue entre P_1 y P_2 y muestra que los problemas (o las situaciones problemáticas) a que el organismo trata de enfrentarse son muchas veces *nuevos* y surgen por sí mismos como resultado de la evolución. Por tanto, la teoría suministra implícitamente una explicación racional de lo que tradicionalmente se ha denominado de modo un tanto equívoco, "*evolución creadora*" o "*evolución emergente*" [55].

(12) Nuestro esquema permite el desarrollo de controles para eliminar errores (órganos de alerta, como los ojos; mecanismos de realimentación); es decir, controles que pueden eliminar los errores sin acabar con el organismo, cosa que hace posible, en última instancia, que nuestra hipótesis muera en lugar nuestro.

XIX

Todo organismo se puede considerar como un sistema jerárquico de controles plásticos —como un sistema de nubes controlado por nubes—. Los subsistemas controlados hacen movimientos de ensayo y error que el sistema de control en parte suprime y en parte restringe.

Ya hemos visto un ejemplo de esta situación en la relación que va de las funciones inferiores del lenguaje a las superiores. Las inferiores siguen existiendo y cumpliendo su papel, aunque se ven constreñidas y controladas por las superiores.

He aquí otro ejemplo típico: Si me mantengo quieto de pie sin hacer ningún movimiento, entonces (según los fisiólogos) mis músculos trabajan constantemente, contrayéndose y relajándose de un modo casi aleatorio (véase TS_1 y TS_n en la tesis (8) de la sección anterior), aunque están siendo controlados, sin que yo me de cuenta, mediante la eliminación de errores (*EE*), de tal manera, que toda desviación de mi posición se corrige casi al instante.

---

[54] La emergencia de una situación problemática nueva podría considerarse como un cambio o diferenciación del "nicho ecológico" o del medio relevante para el organismo. (Tal vez pueda denominarse "selección del hábitat"; cf. B. Lutz, *Evolution, 2*, 1948, págs. 29 y sigs.) El hecho de que *cualquier* cambio en el organismo *o* en sus hábitos *o* en su *hábitat* produzca nuevos problemas explica la increíble riqueza de las soluciones (siempre provisionales).

[55] Véase la nota 24 en relación con las consideraciones de Compton en torno a la "evolución emergente".

De este modo me mantengo en pie sin moverme, aproximadamente de la misma manera que un piloto automático mantiene la estabilidad de la aeronave en su trayectoria.

Este ejemplo ilustra también la tesis (1) de la sección anterior, según la cual todo organismo se enfrenta continuamente a la solución de problemas por ensayo y error y reacciona tanto a los nuevos como a los viejos problemas mediante ensayos más o menos aleatorios [56] o nebulosos que se eliminan si no tienen éxito. (Si lo tienen, aumentan la probabilidad de supervivencia de las mutaciones que "estimulan" las soluciones así obtenidas y tienden a hacerlas hereditarias [57], incorporándolas a la forma o estructura espacial del nuevo organismo.)

XX

Este esquema de la teoría es muy breve y, como es evidente, precisa una gran elaboración. Ahora bien, voy a explicar *uno* de los puntos con más detalle —el uso que hago (en las tesis de la (1) a la (3) de la sección XVIII) de los términos "*problema*", "*resolución de problemas*" y, especialmente, de la afirmación según la cual *podemos hablar de problemas en sentido objetivo y no psicológico.*

Este aspecto es importante porque está claro que la evolución no es un proceso consciente. Muchos biólogos dicen que la evolución de ciertos órganos resuelve determinados problemas; por ejemplo, la evolución del ojo resuelve el problema de avisar a tiempo al animal que se mueve para que cambie de dirección antes de chocar con un obstáculo. Nadie pretende que este tipo de solución a esa clase de problemas sea buscado conscientemente. ¿No es acaso una metáfora hablar de resolución de problemas?

No lo creo; la situación es más bien ésta: cuando hablamos de un problema, casi siempre lo hacemos retrospectivamente. La persona que trabaja en un problema rara vez puede decir con claridad de qué problema se trata (a menos que haya encontrado una solución); y aunque pueda explicar su problema, lo más probable es que se confunda. Es algo que también les ocurre a los científicos a pesar de que ellos son de los pocos que intentan conscientemente alcanzar pleno conocimiento de sus problemas. Por ejemplo, el

---

[56] El método de ensayo y supresión de errores *no opera mediante ensayos totalmente azarosos o aleatorios* (como se ha sugerido alguna vez), aunque dichos ensayos puedan parecer plenamente aleatorios; debe haber al menos una "secuela" [*after-effect*] (en el sentido de mi libro *The Logic of Scientific Discovery*, págs. 162 y siguientes [trad. cit., págs. 151-2].) La razón de ello es que el organismo aprende constantemente de sus errores; esto es, establece controles que suprimen, eliminan o, al menos, reducen la frecuencia de ciertos ensayos *posibles* (que tal vez fuesen *actuales* en una etapa pasada de su evolución).

[57] Algunas veces se llama ahora a esto el "Efecto Baldwin"; véase, por ejemplo, su G. E. Simpson, "The Baldwin Effect", *Evolution*, 7, 1953, págs. 110 y sigs., y C. H. Wadddington, en el mismo volumen, pág. 118 y sigs. (véase especialmente la página 124) y págs. 386 y sig. Véase también J. Mark Baldwin, *Developement and Evolution*, 1902, págs. 174 y sigs., y H. S. Jennings, *The Behaviour of the Lower Organisms*, 1906, págs. 321 y sigs.

problema que Kepler se planteaba conscientemente era el descubrimiento de la armonía del orden del mundo, si bien podemos decir que lo que resolvió fue el problema de la descripción matemática del movimiento de un sistema planetario compuesto por conjuntos de dos cuerpos. De un modo similar, Shrödinger estaba en un error respecto a cuál era el problema que había resuelto con el hallazgo de la ecuación de Schrödinger (independiente del tiempo): creía que sus ondas eran ondas de carga-densidad en un campo continuo de carga eléctrica con intensidad variable. Más tarde, Max Born dio una interpretación estadística de la amplitud de onda de Schrödinger, interpretación que sorprendió al propio Schrödinger y que nunca le llegó a gustar. Había resuelto un problema aunque no el que creía haber resuelto. Pero esto lo sabemos ahora, retrospectivamente.

Sin embargo, es evidente que en la ciencia es donde somos más plenamente conscientes de los problemas que tratamos de resolver. Así pues, no sería inadecuado utilizar también, en otros casos, consideraciones retrospectivas para decir que la ameba resuelve ciertos problemas (aunque no precisamos suponer que sea —en ningún sentido— consciente de sus problemas): de la ameba a Einstein no hay más que un paso.

XXI

Sin embargo, Compton nos dice que los actos de la ameba no son racionales [58], si bien podemos suponer que lo son los de Einstein. Por tanto, después de todo, tiene que haber una diferencia.

Admito que haya una diferencia; aunque sus métodos cuasi-aleatorios y sus movimientos nebulosos de ensayo y error no sean básicamente muy distintos [59, 60], hay una gran diferencia en sus actitudes frente al error. Al contrario que la ameba, Einstein, siempre que se le ocurría una solución nueva, intentaba falsarla conscientemente por todos los medios, detectando en ella algún error: enfocaba *críticamente* sus propias soluciones.

Creo que la diferencia realmente importante que media entre el método de Einstein y el de la ameba es la actitud crítica consciente hacia sus propias ideas. Dicha actitud permitió a Einstein rechazar, rápidamente, cientos de hipótesis inadecuadas antes de pasar a un examen más cuidadoso de alguna de ellas en caso de que pareciese capaz de mantenerse en pie frente a críticas más serias.

Como ha dicho recientemente el físico John Archibald Wheeler. "Nuestro único problema no es otro que lograr que nuestros errores sean lo más

---

[58] Véase *The Freedom of Man*, pág. 91, y *The Human Meaning of Science*, página 73.

[59] Cf. H. S. Jennings, *op. cit.*, págs. 334 y sigs., 394 y sigs. Un bonito ejemplo de resolución de problemas por parte de un pez aparece descrito en el libro de K. Z. Lorenz, *King Salomon's Ring*, 1952, págs. 37 y sig. [traducción castellana del Dr. Ramón Margalef, *El Anillo del Rey Salomón*, Barcelona, Labor, 1962; págs. 60-2].

[60] Vid. nota 56, *supra*.

breves posible" [61]. Este problema de Wheeler se resuelve adoptando conscientemente la actitud crítica. Creo que esta es la forma más elevada que hay de actitud racional o de racionalidad.

Los ensayos y errores de los científicos son hipótesis formuladas verbalmente o, normalmente, por escrito. El científico trata de descubrir fallos en cualquiera de dichas hipótesis mediante la crítica y la contrastación experimental, asistido por sus compañeros científicos que se felicitarán si consiguen descubrir un fallo. Si la hipótesis no se mantiene en pie frente a esas críticas y contrastaciones, al menos con el mismo éxito que sus rivales [62], será eliminada.

Con el hombre primitivo y con la ameba la situación es distinta. Aquí no hay actitud crítica, por lo que lo más normal es que la selección natural elimine una hipótesis o expectativa equivocada, eliminando los organismos que la sostienen o creen en ella. Podemos decir, pues, que el método racional o crítico consiste en dejar que nuestras hipótesis mueran en nuestro lugar: es un caso de evolución exosomática.

XXII

Tal vez pueda abordar ahora una cuestión que me ha dado muchos quebraderos de cabeza, aunque terminé dando con una solución muy sencilla.

La pregunta es: ¿Podemos mostrar la existencia de controles plásticos? ¿Hay en la naturaleza sistemas físicos inorgánicos que suministren ejemplos o modelos físicos de controles plásticos?

Da la impresión de que la pregunta ha recibido implícitamente una respuesta negativa por parte de muchos físicos que, como Descartes o Compton, operan con modelos a base de conmutadores generales, así como por parte de muchos filósofos que, como Hume o Schlick, niegan que pueda existir algo intermedio entre el determinismo completo y el puro azar. Es evidente que los expertos en cibernética y los técnicos en computadoras han logrado construir muy recientemente computadoras a base de elementos férreos [*hardware*] que incorporan controles altamente plásticos; como por ejemplo, las computadoras construidas a base de mecanismos para producir ensayos aleatorios comprobados y evaluados por un mecanismo de retroalimentación (al modo de un piloto automático o un ingenio autorregulado) que los elimina si resultan erróneos. Pero dichos sistemas, aunque contengan lo que he denominado controles plásticos, están compuestos básí-

---

[61] John A. Wheeler, *American Scientist, 44*, 1956, pág. 360.

[62] El que, dado un conjunto de hipótesis rivales, podamos elegir la "mejor" —la "mejor" a la luz de la discusión crítica entregada a la búsqueda de la verdad— significa que, a la luz de la discusión, elegimos la que parece "aproximarse más a la verdad"; véase mi libro *Conjectures and Refutations*, capítulo 10. Véase también *The Freedom of Man*, págs. VII y sigs., especialmente la pág. 74 (en torno al principio de conservación de la energía).

camente de relés complejos de conmutadores generales. Sin embargo, lo que yo andaba buscando era un modelo físico sencillo de indeterminismo peirceano; un sistema físico puro semejante a una nube muy nebulosa en movimiento por calor, controlado por otras nubes nebulosas —aunque tal vez menos nebulosas.

Si volvemos a nuestra vieja ordenación de nubes y relojes, con una nube a la izquierda y un reloj a la derecha, entonces podemos decir que lo que andamos buscando es algo intermedio, como un organismo o una nube de mosquitos, pero que no esté vivo: un sistema físico puro, controlado plástica y "blandamente", como si dijéramos.

Supongamos que la nube a controlar es un gas. Así, podemos situar en el extremo de la derecha un gas incontrolado que pronto se difundirá dejando de constituir un *sistema* físico. En el extremo de la derecha ponemos un cilindro de hierro lleno de gas: es éste nuestro ejemplo de control "duro" o "férreo". En el medio, aunque lejos de la izquierda, hay muchos sistemas controlados de modo más o menos "blando", como el enjambre de mosquitos o las grandes acumulaciones de partículas tales como un gas que se mantiene unido por la gravedad, más o menos como el sol. (No importa que el control sea muy imperfecto y que se escapen muchas partículas.) Podría decirse, tal vez, que los planetas están férreamente controlados por lo que respecta a sus movimientos —relativamente, como es natural, ya que incluso el sistema planetario es una nube como la vía láctea, las aglomeraciones de estrellas y las aglomeraciones de conglomerados. Pero, ¿acaso hay ejemplos de sistemas físicos pequeños controlados "blandamente", a parte de los sistemas orgánicos y esos vastos sistemas de partículas?

Creo que los hay. Propongo que, en medio de nuestro diagrama, coloquemos la pelota de un niño o, tal vez mejor, una burbuja de jabón que resulta ser un ejemplo muy primitivo, y en muchos aspectos excelente, de sistema peirceano *con* un tipo "blando" de control plástico.

La burbuja de jabón consta de dos subsistemas que resultan ser nubes y que se controlan entre sí: sin el aire, la película de jabón se destruiría, con lo que sólo tendríamos una gota de agua jabonosa. Sin la película de jabón el aire estaría descontrolado, se difundiría y dejaría de formar un sistema. Por tanto, el control es mutuo, plástico y tiene un carácter de retroalimentación. Con todo, es posible establecer una distinción entre el sistema controlado (el aire) y el sistema de control (la película): el aire encerrado no sólo es más nebuloso que la película envolvente, sino que además deja de constituir una sistema físico (con interacción) si eliminamos la película. Por el contrario, tras eliminar el aire, la película formará una gota que, aunque tenga una forma distinta, se puede decir que sigue siendo un sistema físico.

Al comparar la burbuja con un sistema "metálico" [*hardware*] como, por ejemplo, un reloj de precisión o una computadora, diremos, naturalmente (de acuerdo con el punto de vista de Peirce), que incluso estos sistemas metálicos son nubes controladas por nubes. Pero estos sistemas "duros" se

construyen a fin de minimizar, en la medida de lo posible, los efectos nebulosos de los movimientos y fluctuaciones de calor molecular: aunque sean nubes, los mecanismos de control están ingeniados para suprimir o compensar todo lo posible los efectos nebulosos. Esto también se aplica a las computadoras provistas de dispositivos que simulan mecanismos aleatorios de ensayo y error.

Nuestra pompa de jabón es distinta y, en este sentido, se asemeja más a un organismo: los efectos moleculares no se eliminan, sino que contribuyen especialmente al funcionamiento del sistema cubierto por una piel —una pared permeable [**] que hace que el sistema sea "abierto", capaz de "reaccionar" a las influencias del medio según pautas incorporadas, como si dijéramos, a su "orgarnización": cuando una radiación térmica se encuentra con la pompa de jabón, ésta absorve el calor (como ocurre en un invernadero) de modo que el aire encerrado se expande y hace que la pompa flote.

No obstante, hemos de buscar los límites de todos estos usos de la semejanza o analogía. Aquí hemos de sañalar que, al menos en algunos organismos, las fluctuaciones moleculares se ven aparentemente amplificadas, dando lugar al desencadenamiento de movimientos de ensayo y error. En todo caso, los amplificadores parecen desempeñar funciones importantes en todos los organismos (que en este sentido se parecen a las computadoras con sus conmutadores generales y cascadas de amplificadores y relés). Sin embargo, en la pompa de jabón no hay amplificadores.

Sea como sea, nuestra pompa muestra la existencia de sistemas físicos nebulosos que están controlados de un modo plástico y blando por otros sistemas nebulosos. (Incidentalmente, como es natural, la película de la pompa no necesita ser obtenida a partir de materia orgánica, aunque habrá de contener moléculas grandes.)

XXIII

La teoría evolucionista, aquí expresada, suministra una solución inmediata a nuestro segundo problema fundamental —el problema cartesiano clásico del cuerpo y la mente—. Se soluciona (sin especificar de *qué* "mente" o "cuerpo" se trata) haciendo unas consideraciones en torno a la evolución y, por tanto, acerca de las funciones de la mente o conciencia.

Hemos de suponer que la conciencia se desarrolla a partir de un pequeño punto de partida; tal vez su primera manifestación sea una vaga sensación de irritación experimentada cuando el organismo ha de resolver un problema como el de sacudirse de encima una sustancia irritante. Sea como

[**] Las paredes permeables o membranas parecen constituir una característica común a todos los sistemas biológicos. (Esto puede ponerse en relación con el fenómeno de la individuación biológica.) Para la prehistoria de la idea de que las membranas y burbujas constituyen organismos primitivos, véase C. H. Kahn, *Anaximander*, 1960, páginas 111 y siguientes.

sea, la conciencia tomará un significado evolucionista —un significado cada vez mayor— cuando empiece a *anticipar* modos posibles de reacción: posibles movimientos de ensayo y error con sus posibles resultados.

Se puede decir que los estados o secuencias de estados conscientes funcionan como sistemas de control, de supresión de errores: la supresión, por regla general, del comportamiento (incipiente) consistente en un movimiento (también incipiente). Desde este punto de vista, la conciencia se muestra como un tipo más de control en interacción. Si pensamos en los sistemas de control incorporados, por ejemplo, en los libros —teorías, sistemas de leyes y todo lo que constituye el "universo del significado"— entonces difícilmente podemos decir que la conciencia es el sistema de control que ocupa el puesto más elevado de la jerarquía, puesto que está en gran medida controlada por estos sistemas lingüísticos exosomáticos, aunque se pueda decir que son un producto de la conciencia. Se puede conjeturar que la conciencia, a su vez, es *producida* por los estados físicos, aunque éstos estén controlados en gran medida por aquella. Del mismo modo que un sistema legal o social nos controla a pesar de ser un producto nuestro y no se puede decir en ningún sentido razonable que sea "semejante" o "igual" a nosotros, aunque entre en interacción con nosotros, así los estados de conciencia (la "mente") controlan el cuerpo y entran en *interacción* con él.

Por tanto, hay todo un conjunto de relaciones de analogía. La conciencia se relaciona con la conducta del organismo individual activo del mismo modo que nuestro mundo de significaciones exosomático se relaciona con la conciencia. Además, el comportamiento del organismo individual se relaciona de un modo semejante con su cuerpo; es decir, con el organismo individual en cuanto sistema fisiológico. Este último se relaciona de manera similar con la sucesión evolutiva de organismos —el *phylum* del que constituye la punta de flecha, como si dijéramos: del mismo modo que el *phylum* lanza experimentalmente el organismo individual a modo de prueba (con lo que éste controla en gran medida el destino del *phylum*), así el sistema fisiológico lanza experimentalmente una conducta del organismo a modo de prueba (con lo que el destino de dicho sistema se ve en gran medida controlado por esa conducta). Nuestros estados de conciencia se relacionan de un modo similar con nuestra conducta, pues la anticipan examinando por ensayo y error sus posibles consecuencias. Por tanto, no sólo controlan, sino que además ensayan *deliberadamente*.

Vemos ahora que esta teoría nos suministra una solución casi trivial del problema de Descartes. Sin decir *qué es "la mente"*, lleva de inmediato a la conclusión de que nuestros *estados mentales controlan (algunos de) nuestros movimientos físicos,* con lo que tiene lugar cierto intercambio, cierta retroalimentación o cierta *interacción* entre la actividad mental y las demás funciones del organismo [64].

---

[64] Como he sugerido en diversos lugares, conjeturo que la aceptación de una "*interacción*" entre los estados físicos y mentales es la única que ofrece una solución satisfactoria al problema de Descartes; véase también la nota 44. Deseo añadir aquí

El control, una vez más, tendrá un carácter "plástico"; de hecho todos sabemos —especialmente los que tocan un instrumento como el piano o el violín— que el cuerpo no siempre hace lo que queremos, así como que hemos de aprender mediante nuestros fracasos a modificar nuestras metas, haciendo concesiones a las limitaciones de nuestros controles: aunque hasta cierto punto seamos libres, siempre hay condiciones físicas o de cualquier otra naturaleza que limitan lo que podemos hacer. (Naturalmente, antes de rendirnos, somos libres de intentar vencer estas limitaciones.)

Por tanto, siguiendo a Descartes, propongo la adopción de un enfoque dualista, aunque como es obvio no propongo establecer una distinción entre *dos tipos de sustancias interactuantes*. Sin embargo, considero útil y legítimo hablar de *dos tipos* de *estados* (o acontecimientos) *en interacción* de carácter físio-químico y mental. También he de señalar que si sólo distinguimos estos dos tipos de estados, entonces continuaremos considerando nuestro mundo desde una perspectiva demasiado estrecha. Debemos distinguir también, por lo menos, los artefactos producidos por los organismos, especialmente los productos de la actividad intelectual, que pueden entrar en interacción con nuestras mentes y, por tanto, con el estado de nuestro medio físico. Aunque frecuentemente dichos artefactos no sean más que "simples porciones de materia", tal vez, "simples herramientas", con todo, a veces constituyen obras de arte consumadas, incluso a nivel animal. A nivel humano, nuestros productos mentales son a menudo mucho más que porciones de materia" (trozos de papel con marcas, por ejemplo), ya que esos trozos de papel pueden representar fases de una discusión o del desarrollo del conocimiento que pueden ir más allá (a veces con consecuencias muy serias) de la capacidad de comprensión de la mayoría de las mentes (si no de todas) que han contribuido a su producción. Por tanto, no podemos conformarnos simplemente con el dualismo, sino que hemos de ser pluralistas y reconocer que las grandes transformaciones que hemos producido en el universo físico, a veces de modo inconsciente, muestran que las ideas y reglas abstractas, algunas de las cuales tal vez la mente humana sólo pueda captar parcialmente, pueden mover montañas.

XXIV

Quisiera señalar un último punto a modo de recapitulación.

Sería un error pensar que la evolución sólo puede llevar, mediante la selección natural, a lo que podríamos denominar resultados "utilitarios": adaptaciones útiles para la supervivencia.

---

que pienso que tenemos buenas razones para suponer que haya estados mentales o conscientes (por ejemplo, los sueños) en los que la conciencia del yo (o de la propia identidad y posición espacio-temporal) es muy débil o inexistente. Por tanto, parece razonable suponer que la conciencia plena del yo es un desarrollo tardío y que resulta equivocado formular el problema del cuerpo y la mente de modo que esta forma de conciencia (o "voluntad" consciente) se considere como la única posible.

Del mismo modo que, en un sistema dotado de controles plásticos, los subsistemas de control y los subsistemas controlados interactúan entre sí, así nuestras soluciones tentativas entran en interacción con nuestros *problemas* y con nuestros *fines*. Esto significa que nuestros fines pueden transformarse y, con ello, *la elección de un fin puede convertirse en algo problemático;* distintos fines pueden entrar en conflicto, dando lugar a la invención de nuevos fines controlados por el método de ensayo y supresión de errores.

Es evidente que si un nuevo fin entra en contradicción con la supervivencia, puede ser eliminado por la selección natural. Es bien sabido que muchas mutaciones son letales y, por tanto, suicidas. Hay muchos casos de fines suicidas, si bien otros pueden ser neutrales para la supervivencia.

Algunos fines que en un principio estaban subordinados a la supervivencia, pueden independizarse más tarde e incluso oponerse a ella. Por ejemplo, la ambición de sobresalir en empresas arriesgadas, escalar el Everest, descubrir un nuevo continente, llegar el primero a la Luna o, también, la ambición de descubrir una nueva verdad.

Otros fines pueden ser autónomos desde el comienzo, independientemente de la supervivencia. Tal vez las tendencias artísticas sean de esta clase, así como los fines religiosos que pueden resultar más importantes que la supervivencia para aquellos que los asumen.

Todo esto forma parte de la superabundancia de la vida —la abundancia casi excesiva de ensayos y errores de que depende el método de ensayo y supresión de errores [••].

Tal vez no carezca de interés ver que tanto los artistas como los científicos emplean efectivamente el método de ensayo y error. Un pintor puede dar tentativamente una pincelada y retroceder para juzgar críticamente el efecto producido [••], a fin de cambiarlo si no resuelve el problema que deseaba resolver. Incluso puede suceder que un efecto accidental o inesperado de su ensayo provisional —una pincelada o una mancha de color— pueda hacerle cambiar de problema, crear un subproblema o una nueva meta: la evolución de los propósitos del artista y de los cánones artísticos (que, como las reglas de la lógica, pueden convertirse en sistemas de control exosomático) avanza también por el método de ensayo y error.

Tal vez podamos detenernos aquí un momento para volver sobre el problema del determinismo físico y sobre nuestro ejemplo del físico sordo que, sin haber oído nunca música, era capaz de "componer" una ópera de Mozart o una Sinfonía de Beethoven, sólo con estudiar sus cuerpos y el medio en que se desenvolvían sus sistemas físicos, prediciendo en qué lugar del pentagrama iban a escribir sus plumas las señales negras. Para mí, estas

---

[••] Cf., por ejemplo, mi libro *Conjectures and Refutations*, especialmente la página 312 [trad. cit., pág. 359].

[••] Véase, por ejemplo, Ernst H. Gombrich, *Meditations on a Hobby Horse*, especialmente la pág. 10 [traducción castellana, *Meditaciones sobre un Caballo de Juguete*, Barcelona, Seix-Barral, 1968], así como el libro del mismo autor, *Art and Illusion*, 1960, 1962 (véase en el índice la expresión "ensayo y error"). Cf. también la nota 67.

consecuencias del determinismo físico son inaceptables. Mozart y Beethoven están controlados, en parte, por sus "gustos", por sus sistemas de evaluación musical. Ahora bien, dichos sistemas no son férreos, sino plásticos. Responden a nuevas ideas y se pueden modificar mediante nuevos ensayos y errores —tal vez, incluso, mediante un error accidental o una disonancia inesperada [97].

Resumiré la situación para terminar.

Hemos visto que no resula satisfactorio considerar el mundo como un sistema físico cerrado —sea un sistema extrictamente determinista o un sistema en que lo que no está extrictamente determinado se deba sencillamente al azar: si consideramos el mundo de este modo, la creatividad y libertad humanas no puede ser mas que ilusiones. El intento de recurrir a la indeterminación de la teoría cuántica también es insatisfactorio, pues nos conduce al azar más bien que a la libertad y a decisiones precipitadas más bien que a decisiones deliberadas.

Por tanto, he ofrecido aquí una visión del mundo distinta (según la cual el mundo físico es un sistema abierto) compatible con la consideración de la evolución de la vida como un proceso de ensayo y supresión de errores. Además, nos permite comprender racionalmente, aunque no plenamente, la emergencia de novedades biológicas, así como el desarrollo del conocimiento y libertad humanos.

He intentado bosquejar una teoría evolucionista que dé cuenta de todo esto y que ofrezca soluciones a los problemas de Compton y Descartes. Temo que resulte una teoría demasiado farragosa y especulativa. Aunque creo que de ella se pueden sacar consecuencias contrastables, no creo, ni mucho menos, que sus conclusiones sean las que andaban buscando los filósofos. Sin embargo, creo que Compton habría considerado que, a pesar de sus fallos, suministra una respuesta posible y que, además, puede conducir a un avance ulterior.

---

[97] Para la estrecha semejanza existente entre los productos artísticos y científicos, véase *The Freedom of Man*, Prólogo, págs. VII y sigs., y las consideraciones que aparecen en la pág. 79 a las que hemos aludido en la nota 62, *supra*. Véase también E. Mach *Wärmelehre*, 1896, págs. 440 y sig. donde escribe: "La historia del arte ...nos enseña que en las obras de arte se utilizan formas que surgen accidentalmente. Leonardo da Vinci recomienda al artista la búsqueda de formas en las nubes o de manchas en las paredes sucias y ahumadas que puedan sugerirle ideas concordes con sus planes y estilo... Así mismo, un músico puede a veces sacar ideas de ruidos aleatorios. Incluso, a veces, oímos decir a un famoso compositor que ha descubierto motivos armónicos o melódicos valiosos tras haber dado accidentalmente una mala nota cuando tocaban el piano".

7. LA EVOLUCION Y EL ARBOL DEL CONOCIMIENTO*

Es para mí muy grato haber sido invitado a pronunciar la conferencia Herbert Spencer y no sólo por el honor que representa haber sido llamado para rendir homenaje a un pensador de gran audacia y originalidad. Lo que me agradó especialmente fue la sugerencia, hecha por el Comité Organizador de estas Conferencias, en el sentido de que debería elegir un tema en torno a "El Método de las Ciencias Biológicas". Dicha sugerencia me brinda la oportunidad de desarrollar aquí determinadas ideas que, aunque pienso que son estimulantes y dignas de discusión, nunca habría presentado en público de no haber sido por este estímulo.

Todas las ideas que voy a exponer ante ustedes versan sobre problemas de método de la Biología. Con todo, no me limitaré a dicho campo. Mi plan consiste en dividir esta conferencia en tres apartados. Comenzaré haciendo algunas consideraciones en torno a la teoría general del conocimiento, volveré luego sobre determinados problemas de método importantes para la teoría de la evolución para entrar o, más bien, chapotear en determinados aspectos de la propia teoría de la evolución. Concretamente, en la tercera parte de mi conferencia les expondré a ustedes una conjetura mía que pretende resolver, desde el marco de la teoría darwinista o neodarwinista de la selección natural, algunas dificultades clásicas en las que se ha debatido hasta ahora la teoría.

Digo que estas dificultades son "clásicas" porque fueron descubiertas en seguida y analizadas sucintamente, tanto por parte de Herbert Spencer, al poco tiempo de aceptar la teoría darwinista de la selección natural, como por parte de Samuel Butler, al poco de rechazarla. Por otra parte, como señaló Spencer, el propio Darwin se había ocupado de las dificultades a que me refiero [1].

---

* Este artículo se basa en la Conferencia Herbert Spencer, pronunciada en Oxford el 30 de octubre de 1961. Los añadidos más importantes, así como las notas totalmente nuevas, aparecen entre corchetes. El Apéndice ha sido añadido en 1971.

[1] Véase especialmente el artículo de Spencer "The Factors of Organic Evolution". publicado por primera vez en sus *Essays* (por ejemplo, en el vol. I de la "Library Edition" de 1891, págs. 389 y sigs.). Es interesante señalar que entre las muchas ideas importantes que contiene dicho artículo se formula un enfoque que hoy día se denomina "enfoque organísmico de la biología" y que es tenido por una innovación; véase, por ejemplo, la pág. 410, donde Spencer habla de cambios en determinados órganos diciendo que "todos los demás ...órganos se ven implicados en este cambio. Las funciones que realizan han de constituir *un equilibrio dinámico*". (El subrayado es mío.) Spencer describe aquí el organismo en términos actuales como "un sistema abierto en equilibrio fluido" (o en "un estado aproximadamente estable").

El programa de mi conferencia abarca, por tanto, desde la teoría general del conocimiento hasta la misma teoría de la evolución, pasando por los métodos de la biología. Me temo que este programa sea un tanto ambicioso para una conferencia. Si, además, mi programa incluyese la pretensión de *convencerles* a ustedes, entonces mi posición será desesperada. Por tanto, es una suerte que no pretenda convencer a nadie de la verdad de ninguna de mis tesis y, menos que nada, de la verdad de mi conjetura neo-darwinista que propondré al final de mi conferencia. Aunque abrigo la esperanza de que tal vez dicha conjetura nos ayude a aproximarnos un tanto a la verdad, no tengo la osadía ni tan siquiera de confiar en que sea verdadera. Por otra parte, me temo que contenga muy poco de verdad. Ciertamente, no representa la verdad definitiva ni toda la verdad sobre el tema. Por tanto, si no pretendo convercerles a ustedes, ello se debe a la sencilla razón de que ni yo mismo estoy convencido. Sin embargo, deseo —y pondré todos los medios para ello— despertar su interés por estos problemas. Admito que estos problemas se han convertido a veces en problemas un tanto anticuados, e incluso he expresado en alguna parte* mi acuerdo con la consideración del profesor Raven según la cual la controversia evolucionista no es sino "una tormenta en una taza de té victoriana". Si bien esta descripción encaja perfectamente con la tormenta desencadenada por la tesis de Darwin relativa a nuestro parentesco con los monos, ha habido otros problemas teóricos, para mí mucho más estimulantes, planteados por la controversia darwinista.

1. Algunas consideraciones en torno a los problemas y desarrollo del conocimiento

Vamos ahora con la primera parte de mi conferencia: la teoría general del conocimiento.

La razón por la cual pienso que debo empezar con algunos comentarios en torno a la teoría del conocimiento es que estoy en desacuerdo con casi todo el mundo a este respecto, si exceptuamos quizá a Charles Darwin y a Albert Einstein. (Einstein, dicho sea de paso, expuso su punto de vista sobre estas cuestiones en su conferencia Herbert Spencer de 1933.) [2] El punto fundamental es la relación entre observación y teoría. Creo que la teoría siempre va delante —al menos una teoría o expectativa rudimentaria—;

---

* Cf. *La miseria del Historicismo*, trad. cast. de Pedro Schwartz, Madrid, Taurus, 1961, pág. 132, nota 2. (Hay reedición en Alianza Editorial, Madrid, 1973.) Véase también la nota 52 del capítulo 6 del presente volumen. [*N. T.*]

[2] Albert Einstein, *On the Methods of Theoretical Physics*, 1933. (También en *The World As I See it.*) ["Sobre el Método de la Física Teórica" en *Mi Perspectiva Mundial*, traducción castellana de Nathan Caplán, Buenos Aires, Santiago Rueda Editor, 1946, págs. 159-169.] Sir Peter Medawar me ha indicado que además de a Darwin y a Einstein debería citar aquí a Claude Bernard, *An Introduction to the Study of Experimental Medicine* (1865), 1927. [*Introducción al estudio de la Medicina Experimental*, traducción castellana de Luis Alberti, Ediciones El Centaurio, Madrid, 1947.]

siempre precede a las observaciones cuyo papel fundamental, así como el de las contrastaciones experimentales, es mostrar que algunas de nuestras teorías son falsas, estimulándonos de este modo a construir otras mejores.

Por consiguiente, afirmo que no partimos de observaciones, sino siempre de *problemas* —sea de problemas prácticos o de *una teoría que se encuentra en dificultades*—. Una vez que abordamos un problema podemos comenzar a trabajar sobre él, lo cual se puede hacer en dos pasos: podemos proceder tratando, en primer lugar, de tramar o conjeturar una solución al problema, para luego intentar criticar nuestra suposición que, normalmente, será más bien endeble. A veces, nuestra suposición o conjetura podrá mantenerse durante algún tiempo frente a nuestra crítica y nuestras contrastaciones experimentales. Mas, por regla general, pronto encontraremos que nuestras conjeturas se pueden refutar, que no resuelven nuestro problema o que sólo lo resuelven en parte. Además, nos encontramos con que incluso las mejores soluciones —las que son capaces de resistir las críticas más rigurosas de las inteligencias más brillantes o ingeniosas— en seguida plantean dificultades y problemas. Por tanto, podemos decir que *el aumento del conocimiento va de viejos a nuevos problemas mediante conjeturas y refutaciones.*

Supongo que algunos de ustedes estarán de acuerdo en que generalmente partimos de un problema. Sin embargo, es posible que sigan pensando que nuestros problemas tienen que haber surgido de observaciones y experimentos, puesto que todos ustedes están familiarizados con la idea de que nada puede haber en nuestro intelecto que no hayamos recibido a través de los sentidos.

Ahora bien, lo que yo ataco es precisamente esta idea venerable [3]. Afirmo que todo animal ha nacido con expectativas o anticipaciones que pueden tomarse como hipótesis; una especie de conocimiento hipotético. Afirmo, además, que en este sentido poseemos un determinado grado de conocimiento innato del cual partir, aunque sea poco fiable. Este conocimiento innato, estas expectativas innatas crearán *nuestros primeros problemas,* si se ven defraudadas. Podemos decir, por tanto, que el ulterior desarrollo del conocimiento consistirá en corregir y modificar el conocimiento previo.

Por tanto, invierto los términos de quienes piensan que la observación debe preceder a las expectativas y problemas. Incluso afirmo que la observación no puede, por *razones lógicas,* ser anterior a algunos problemas —por ejemplo, a los problemas que surgen de una observación que frusta una expectativa o refuta una teoría—. El que las observaciones no puedan ser previas a los problemas se puede ilustrar mediante un experimento sencillo que, si me permiten, voy a realizar utilizándolos a ustedes como sujetos de ex-

---

[3] [He intentado retrotraer esta doctrina hasta Parménides, quien la formuló para atacarla. Véase la pág. 165 de la segunda edición (1965) de mi libro *Conjectures and Refutations.* [Traducción castellana de Néstor Míguez. *El Desarrollo del Conocimiento Científico. Conjeturas y Refutaciones.* Buenos Aires, Paidos, 1967; págs. 192 y sig.]]

perimentación [4]. Mi experimento consiste en pedirles que *observen* aquí y ahora. ¡Espero que todos ustedes cooperen y observen! Pero me temo que algunos de ustedes, en vez de observar, sentirán unas ganas inmensas de preguntar: "¿QUÉ es lo que quiere usted que observe?".

Si es ésta su respuesta, entonces mi experimento ha tenido éxito, pues lo que trato de ilustrar es que, para observar, hemos de tener *in mente* un problema concreto que podamos zanjar por observación. Darwin sabía esto cuando escribió: "Es extraño que haya quien no vea que toda observación tiene que estar en favor o en contra de algún punto de vista..." [5]. [Si decimos "¡observe!" (sin indicar *el qué*) u "¡observe esta araña!", no estamos dando una orden clara, aunque sí lo haremos si decimos "¡Observe *si* esta araña sube o baja, tal como yo supongo!"]

Naturalmente, no espero convencerles a ustedes de la verdad de mi tesis según la cual las observaciones vienen después de las expectativas o hipótesis. Mas espero haberles podido mostrar que puede haber una alternativa a la doctrina venerable según la cual el conocimiento, especialmente el científico, parte siempre de observaciones [6].

Consideremos ahora con mayor detenimiento este método de conjeturas y refutaciones que, según mi tesis, es el método mediante el que se desarrolla nuestro conocimiento.

Comenzamos con un problema o una dificultad que puede ser práctica o teórica. Sea como sea, cuando nos enfrentamos por vez primera a un problema, es obvio que no podemos saber mucho de él. A lo sumo, sólo tenemos una idea vaga de en qué consiste realmente nuestro problema. ¿Cómo podemos, entonces, dar una solución adecuada? Es obvio que nos resulta imposible. Primero hemos de familiarizarnos más con el problema. ¿Pero de qué modo?

Mi respuesta es muy simple: inventando una solución inadecuada y *criticándola*. Sólo de este modo podemos comprender el problema, pues comprender un problema significa comprender sus dificultades y comprender sus dificultades significa comprender por qué no es fácil de resolver, por qué no marchan las soluciones más obvias. Hemos de inventar, por tanto, esas soluciones más obvias y criticarlas para descubrir *por qué* no marchan. De este modo, nos familiarizamos con el problema y pasamos de las malas soluciones a otras mejores —siempre que supongamos que disponemos de una capacidad creadora para inventar suposiciones una y otra vez.

---

[4] [Doy aquí cuenta de un experimento ya descrito en la pág. 46 de la segunda edición (1965) de mi libro *Conjectures and Refutations* [trad. cit., pág. 58].]

[5] Francis Darwin (ed.), *More Letters of Charles Darwin*, vol. I, 1903, pág. 195. Véase también J. O. Wisdom, *Foundations of Inference in Natural Science*, 1952, página 50 y Nora Barlow, *The Autobiography of Charles Darwin*, 1958, pág. 161. El pasaje de Darwin finaliza con las palabras (que admito que debilitan el apoyo que dicho pasaje presta a mi tesis) "¡Si ha de servir para algo!".

[6] La doctrina aún más venerable según la cual todo conocimiento parte de percepciones o sensaciones que, naturalmente, también rechazamos aquí, está a la base del hecho de que los "problemas de percepción" sigan siendo considerados como una parte respetable de la filosofía o, más exactamente, de la teoría del conocimiento.

Creo que es a esto a lo que nos referimos cuando hablamos de "trabajar sobre un problema". Si hemos trabajado sobre él durante un tiempo suficiente con la necesaria intensidad, empezamos a conocerlo y a comprenderlo en el sentido de saber qué tipo de suposición, conjetura o hipótesis no vale para nada, porque no afecta sencillamente al problema, y qué tipo de requisito debe reunir un intento serio de resolverlo. En otras palabras, empezamos a ver las ramificaciones del problema, sus sub-problemas y su conexión con otros problemas. (Sólo a este nivel se puede someter una solución conjeturada a la crítica de los demás e incluso publicarla.)

Si tomamos ahora en consideración este análisis, descubrimos que encaja con nuestra fórmula según la cual el conocimiento progresa pasando de viejos problemas a otros nuevos, mediante conjeturas e intentos críticos de refutarlas. Incluso el proceso de familiarizarse más y más con un problema procede según esta fórmula. El paso siguiente consiste en discutir y criticar nuestra solución provisional; todo el mundo trata de encontrarle un fallo y refutarla y, sea cual sea el resultado de estos intentos, es evidente que algo nos enseñarán. Si la crítica de nuestros amigos u oponentes tiene éxito, habremos aprendido mucho acerca de nuestro problema: sabremos más que antes acerca de sus dificultades inherentes. Aún cuando no tengan éxito nuestros críticos más penetrantes y nuestra hipótesis logre resistir sus críticas, aún entonces habremos aprendido mucho, tanto sobre el problema como sobre nuestra hipótesis, su adecuación y sus ramificaciones. En la medida en que nuestra hipótesis sobreviva o, al menos, en la medida en que se comporte frente a la crítica mejor que sus rivales, podrá pasar a formar parte, temporal y provisionalmente, de la enseñanza ordinaria de la ciencia.

Podemos decir todo esto señalando que el aumento de nuestro conocimiento es el resultado de un proceso muy similar a lo que Darwin llamaba "selección natural"; es decir, *la selección natural de hipótesis*: nuestro conocimiento consta en todo momento de aquellas hipótesis que han mostrado su aptitud (comparativa), habiendo sobrevivido hasta el momento actual en su lucha por la existencia; lucha competitiva que elimina aquellas hipótesis inadecuadas [7].

Esta interpretación se puede aplicar al conocimiento animal, al conocimiento precientífico y al científico. Lo pecualiar del conocimiento científico es que en él la lucha por la existencia se hace más dura mediante la crítica consciente y sistemática de nuestras teorías. Así pues, mientras que el conocimiento animal y el precientífico se desarrollan fundamentalmente mediante la supresión de quienes sostienen hipótesis inadecuadas, la crítica científica logra frecuentemente que nuestras teorías perezcan en lugar nues-

---

[7] Véase mi libro *The Logic of Scientific Discovery*, especialmente las págs. 108 y 131 [traducción castellana de Víctor Sánchez de Zavala, *La Lógica de la Investigación Científica*, Madrid, Tecnos, 1962, págs. 102 y sigs.], así como mi libro *Poverty of Historicism*, pág. 133 [traducción castellana de Pedro Schwartz, *La Miseria del Historicismo*, Madrid, Taurus, 1961, págs. 163 y sig.]

tro, eliminando así nuestras creencias equivocadas antes de que éstas nos lleven a nuestra propia perdición.

Al plantear así la situación, pretendo describir cómo se desarrolla realmente el conocimiento. No es una metáfora, aunque sea obvio que se utilizan metáforas. La teoría del conocimiento que deseo proponer es una teoría del desarrollo del conocimiento en gran medida darwinista. De la ameba a Einstein, el desarrollo del conocimiento es siempre el mismo: intentamos resolver nuestros problemas, así como obtener, mediante un proceso de eliminación, algo que se aproxime a la adecuación en nuestras soluciones provisionales.

Y, con todo, algo nuevo ha surgido al nivel humano. Para hacerlo ver, compararé el árbol de la evolución con lo que podríamos denominar el creciente árbol del conocimiento.

El árbol de la evolución crece desarrollando cada vez más ramas a partir de un tronco común. Es como un árbol ordinario: el tronco común está formado por nuestros antecesores unicelulares comunes, los antecesores de todos los organismos. Las ramas representan desarrollos tardíos, muchos de los cuales se han "diferendiado" —para decirlo con la terminología de Spencer— en formas altamente especializadas, cada una de las cuales está hasta tal punto "integrada" que puede resolver sus dificultades particulares, sus problemas de supervivencia.

El árbol de la evolución de nuestras herramientas e instrumentos ofrece un aspecto muy similar. Es de presumir que haya empezado a partir de una piedra y un palo, pero bajo el influjo de problemas cada vez más especializados se ha ramificado en un inmenso número de formas altamente especializadas.

Ahora bien, si comparamos ahora estos árboles evolucionistas en desarrollo con la *estructura de nuestro conocimiento en desarrollo,* nos encontramos con que el árbol del conocimiento humano en crecimiento posee una estructura manifiestamente distinta. Está claro que el desarrollo del conocimiento aplicado es muy similar al desarrollo de herramientas y otros instrumentos: siempre constituyen aplicaciones cada vez más diversas y especializadas. Mas el conocimiento puro (o "investigación fundamental" como se le llama a veces) se desarrolla de un modo muy distinto. Se desarrolla casi en sentido opuesto a esta especialización y diferenciación progresiva. Como señaló Herbert Spencer, está dominado en gran medida por la tendencia hacia una integración creciente, hacia teorías unificadas [8]. Esta tendencia se hizo patente cuando Newton combinó la mecánica terrestre de Galileo con la teoría kepleriana de los movimientos celeste. Desde entonces, ha perdurado.

Cuando hablábamos del árbol de la evolución, suponíamos, como es

[8] [Spencer también escribe, criticando a Comte (*Essays,* 1891, vol. II, pág. 24): "El progreso de la ciencia es doble. Procede a la vez de lo particular a lo general y de lo general a lo particular. Es al mismo tiempo analítico y sintético". Spencer cita, como ejemplo de este principio, diez descubrimientos físicos entre los que se cuentan las teorías de Galileo y Newton (ibid., págs. 25 y sigs.).]

obvio, que la dirección del tiempo señalaba hacia arriba —la dirección en que crece el árbol—. Suponiendo la misma dirección del tiempo, habremos de representar el árbol del conocimiento como surgiendo de incontables raíces que crecen en el aire, más bien que bajo tierra, y que, finalmente, tienden a unirse en un tronco común. En otras palabras, la estructura evolucionista del desarrollo del conocimiento puro es casi la opuesta a la del árbol de la evolución de los organismos vivos, los instrumentos humanos o el conocimiento aplicado.

Hemos de explicar ahora este crecimiento integrador del árbol del conocimiento puro. Es el resultado del objetivo peculiar que nos proponemos en la búsqueda del conocimiento puro, consistente en satisfacer nuestra curiosidad explicando las cosas. Es también el resultado de la existencia del lenguaje humano que no sólo nos permite describir el estado de las cosas, sino también argumentar acerca de la verdad de nuestras descripciones; es decir, criticarlas.

Al buscar el conocimiento puro nos proponemos como meta, sencillamente, comprender y resolver el cómo y el por qué, cosa que conseguimos mediante la explicación. Por tanto, todos los problemas del conocimiento puro son *problemas de explicación.*

Dichos problemas pueden tener su origen en problemas prácticos. Así, el problema práctico, "¿Qué se puede hacer para combatir la pobreza?" ha desembocado en el problema puramente teórico, "¿Por qué hay pobres?" y de aquí hemos pasado a la teoría de los salarios y los precios, etc.; en otras palabras, a la teoría económica pura que, como es natural, crea continuamente sus propios problemas. En este desarrollo, los problemas tratados —especialmente los que no logramos resolver— se multiplican y diferencian como ocurre siempre que nuestro conocimiento crece. Con todo, *la misma teoría explicativa* ha mostrado ese crecimiento integrador que Herbert Spencer fue el primero en describir.

Para poner otro ejemplo análogo de carácter biológico, nos enfrentamos al urgentísimo problema práctico de combatir epidemias como la de la viruela. Sin embargo, de la práctica de la inmunización pasamos a la teoría de la inmunología y de ésta a la teoría de la creación de anticuerpos, campo biológico famoso por la profundidad de sus problemas y por la capacidad de multiplicación que éstos poseen.

Los problemas de explicación se resuelven proponiendo teorías explicativas que pueden ser sometidas a crítica mostrando que, o bien son en sí mismas inconsistentes, o bien son incompatibles con los hechos o con cualquier otro tipo de conocimiento. Ahora bien, esta crítica da por sentado que deseamos dar con teorías *verdaderas* —teorías que concuerden con los hechos—. Creo que es esta idea de *verdad como correspondencia con los hechos* la que hace posible la crítica racional. El crecimiento integrador del árbol del conocimiento se explica gracias a nuestro objetivo de aproximación a la verdad, junto con el hecho de que nuestra curiosidad, nuestra pasión de explicar mediante teorías unificadas, es universal e ilimitada.

Al señalar la diferencia que hay entre el árbol de la evolución de los instrumentos y el del conocimiento puro, deseo ofrecer, incidentalmente, algo así como una refutación del punto de vista, tan de moda, según el cual el conocimiento humano sólo puede entenderse como un instrumento al servicio de la lucha por la supervivencia. Esto ha de servir como advertencia contra una interpretación excesivamente estrecha de lo que he dicho acerca del método de conjeturas y refutaciones y acerca de la supervivencia de la hipótesis más apta. Sin embargo, no entra en conflicto con lo que he dicho, porque no he afirmado que la hipótesis más apta sea siempre la que favorezca nuestra propia supervivencia. He dicho, más bien, que la hipótesis más apta es la que resuelve mejor el *problema* en vista del cual se inventó, así como aquella que resiste las críticas mejor que las hipótesis rivales. Si nuestro problema es puramente teórico —consistente en encontrar una explicación puramente teórica— entonces la crítica se guiará por la idea de aproximación a la verdad, más bien que por la idea de favorecer nuestra supervivencia.

Al hablar aquí de verdad, quiero aclarar que nuestro objetivo consiste en dar con teorías verdaderas o, al menos, con teorías que se aproximen más a la verdad que las conocidas hasta el presente. No obstante, esto no quiere decir que podamos saber con certeza si nuestras teorías explicativas son verdaderas. Lo que podemos hacer es criticar una teoría explicativa y demostrar su falsedad. Mas, una buena teoría explicativa es siempre una anticipación audaz de lo que vendrá. Ha de ser contrastable y criticable, pero no podremos mostrar que sea verdadera; y si tomamos la palabra "probable" en alguno de los diversos sentidos que satisfacen el cálculo de probabilidades, entonces nunca podremos mostrar que es "probable" (es decir, más probable que su negación).

Este hecho no es en absoluto sorprendente, ya que nada ha cambiado aunque hayamos adquirido el arte de la crítica racional y la idea reguladora de que la explicación verdadera es la que corresponde a los hechos. El procedimiento fundamental para el desarrollo del conocimiento sigue siendo el de conjeturas y refutaciones, el de la supresión de explicaciones inadecuadas; y puesto que la eliminación de un número finito de tales explicaciones no puede reducir el número infinito de explicaciones posibles supervivientes, tanto Einstein como la ameba pueden errar.

Por tanto, no podemos atribuir la verdad o la probabilidad a nuestras teorías. La utilización de normas tales como la verdad y la aproximación a la verdad sólo desempeñan un papel en el contexto de la crítica. Podemos rechazar una teoría por no ser verdadera o por aproximarse menos a la verdad que una de sus predecesoras o competidoras.

Tal vez pueda reunir en dos tesis breves lo que he dicho hasta aquí.

(i) Somos falibles y propendemos al error; sin embargo, podemos aprender de nuestros fallos.

(ii) No podemos justificar nuestras teorías, pero podemos criticarlas

racionalmente y adoptar de modo provisional aquellas que parecen soportar mejor nuestras críticas y que poseen el mayor poder explicativo.

Con esto termina la primera parte de mi conferencia.

2. Consideraciones en torno a los métodos de la biología y en especial de la teoría de la evolución

En la segunda parte de mi conferencia —que he tenido que recortar drásticamente para dejar sitio a la tercera parte— me propongo discutir brevemente un cierto número de problemas relativos a los métodos de la biología.

Empezaré con dos tesis generales. He aquí la primera:

(1) Quien piense que el método científico es un medio para tener éxito en ciencia, se verá defraudado. No hay vía real para el éxito.

He aquí mi segunda tesis:

(2) Quien piense que el método científico —o *El Método Científico*— es un modo de justificar los resultados científicos, también se verá defraudado. Un resultado científico no se puede justificar, sino tan sólo criticar y contrastar. Lo más que se puede decir en su favor es que, tras todas las críticas y contrastaciones, parece mejor, más interesante, más potente, más prometedor y más próximo a la verdad que sus rivales.

A pesar de estas dos tesis intencionalmente descorazonadoras, se puede decir algo más positivo. Hay algo así como un secreto del éxito y lo voy a revelar. Es el siguiente:

En todos los estadios de las investigaciones, tened lo más claro posible cuál es vuestro problema y observad de qué modo cambia y se hace cada vez más definido. Tened lo más claro posible cuáles son las diversas teorías que aceptáis y tened presente que todos aceptamos teorías inconscientemente o las damos por supuestas, a pesar de que la mayoría de ellas son falsas casi con certeza. Intentad, una y otra vez, formular explícitamente las teorías aceptadas para criticarlas. Tratad también de construir teorías alternativas —incluso de aquellas teorías que os parezcan ineludibles—, pues sólo así podréis comprender las teorías aceptadas. Siempre que una teoría os parezca la única posible, interpretad esta situación como un síntoma de que ni habéis comprendido la teoría ni el problema que ésta pretende resolver. Considerad siempre los experimentos como contrastaciones de la teoría —como intentos de encontrarles fallos y rechazarlas—. Si un experimento u obserbación parece apoyar una teoría, recordad que lo único que hace es debilitar alguna teoría rival —tal vez una en la que nunca habéis pensado antes—. Que sea vuestra ambición refutar y sustituir vuestras mejores teorías: es mejor que defenderlas, dejando a otros la tarea de refutarlas. Pero

recordad también que una buena defensa de una teoría forma parte, necesariamente, de la discusión fructífera, pues sólo defendiéndola podemos descubrir su fuerza, así como la fuerza de la crítica dirigida contra ella. No hay modo de discutir o criticar una teoría si no intentamos continuamente formularla de la manera más potente, argumentando contra ella sólo por respecto a dicha formulación.

Utilizando una distinción formulada y empleada por Sir Peter Medawar en sus Conferencias Reith [9], podemos decir que el proceso de descubrimiento o de aprendizaje acerca del mundo que he descrito aquí es *evocativo* más bien que *instructivo*. Aprendemos del medio no porque él nos instruya, sino porque nos lanza un desafío: evoca nuestras respuestas (entre ellas, nuestras expectativas, anticipaciones y conjeturas) y aprendemos mediante la supresión de las respuestas estériles —es decir, *aprendemos de nuestros errores*—. No obstante, un método evocativo como éste puede *simular o imitar* un proceso de instrucción: el resultado del proceso puede dar la impresión de que hemos obtenido nuestras teorías partiendo de la observación y procediendo por inducción. La idea de un proceso evocativo de evolución que *simula* un proceso de instrucción es típica de Darwin y desempeña un papel importante en lo que sigue.

El descubrimiento darwinista de la teoría de la selección natural se ha comparado muchas veces con el descubrimiento newtoniano de la teoría de la gravitación. Es un error. Newton formuló un conjunto de leyes universales con el fin de describir la interacción y comportamiento consiguiente del universario físico. La teoría de la evolución de Darwin no propuso tales leyes universales. No hay leyes darwinistas de evolución. De hecho, fue Herbert Spencer quien trató de formular las leyes universales de la evolución —las leyes de "diferenciación" e "integración"—. Como he intentado señalar, dichas leyes no carecen de interés e incluso pueden ser verdaderas. Mas son vagas y, comparadas con las leyes de Newton, están casi vacías de contenido empírico. (El propio Darwin encontraba que las leyes de Spencer poseían un interés más bien escaso.)

No obstante, la influencia revolucionaria de Darwin sobre nuestra imagen del mundo en torno fue, por lo menos, tan grande —si bien no tan profunda— como la de Newton, ya que la teoría darwinista de la selección natural mostró que *en principio* era *posible reducir la teleología a la causación, mediante la explicación, en términos puramente físicos, de la existencia en el mundo de planes y propósitos.*

Lo que Darwin nos mostró fue que el mecanismo de selección natural puede simular, en principio, los actos del Creador junto con sus planes y designios, del mismo modo que puede simular la actividad humana racional orientada a un objetivo o finalidad.

Si esto fuese correcto, podríamos decir, desde el punto de vista del *mé-*

---

[9] [Peter B. Medawar, *The Future of Man*, Methuen, 1961.]

todo biológico: Darwin mostró que tenemos plena libertad para recurrir en biología a explicaciones teleológicas —incluso quienes creemos que toda explicación debe ser causal—. La razón de ello estriba precisamente en que lo que él mostró es que, *en principio* toda explicación teleológica podrá algún día ser reducida a —o ulteriormente explicada en términos de— una explicación causal.

Si bien esto constituye un logro inmenso, hemos de añadir que la expresión *en principio* constituye una restricción importantísima. Hasta el presente, ni Darwin ni ningún darwinista ha suministrado una explicación causal de la evolución adaptativa de un sólo organismo o de un sólo órgano. Lo único que se ha mostrado —y no es poca cosa— es que tales explicaciones pueden existir (es decir, no son lógicamente imposibles).

Ni que decir tiene que mi punto de vista sobre el darwinismo puede ser muy objetable para la mayoría de los biólogos que crean que las explicaciones teleológicas en biología son tan rechazables, o casi, como las teológicas. Su influencia fue lo bastante fuerte para obligar a un hombre como Sir Charles Sherrington a sostener, de un modo defensivo que "no obtendremos el provecho debido del estudio de un tipo de reflejo particular, si no podemos discutir su propósito inmediato considerándolo como un acto adaptado" [10].

Uno de los puntos más obvios del darwinismo —que es importante para la tercera parte de mi conferencia— es que sólo es probable que sobreviva, de hecho, el organismo que muestre en su comportamiento una tendencia, disposición o propensión fuerte a luchar por su supervivencia. En consecuencia, tal disposición tenderá a pasar a formar parte de la estructura genética de todo organismo; aparecerá en su comportamiento y en la mayor parte de su organización —si no en toda ella—. Ciertamente, esto equivale no ya a simular simplemente, sino también a explicar la teleología por medio de la selección natural, aunque sólo sea en principio.

De modo similar podemos decir que el lamarckismo, especialmente la doctrina según la cual los órganos evolucionan *bajo la influencia de su uso* y degeneran bajo el influjo de su desuso, se ha visto *explicado*, en cierto sentido, en términos de selección natural gracias a J. M. Baldwin (un filósofo de Princeton) [a Waddigton, a Simpson] y a Erwin Schrödinger [11]. No analizaré aquí su método de explicación, ya que lo desarrollaré más detenidamente y lo ampliaré, creo, de un modo considerable en la hipótesis que expondré en la tercera parte. Mas quiero dejar bien sentado que lo que Baldwin [Waddington, Simpson] y Schrödinger han mostrado es de qué

---

[10] Cito el libro de Sir Charles Sherrington, *The Integrative Action of the Nervous System*, 1906, 1947, págs. 238.

[11] Véase J. M. Baldwin, *Development and Evolution*, 1902, y Erwin Schödinger, *Mind and Matter*, 1958, especialmente el capítulo "Feigned Lamarckism", págs. 26 y siguientes. [Antes también aludía aquí al libro de Sir Julian Huxley, *Evolution-The Modern Synthesis*, 1942. Sir Peter Medawar ha llamado mi atención sobre el hecho de que la referencia es dudosa en este contexto, así como sobre el artículo de Waddington; véase la nota 57 del capítulo 6.]

modo la evolución lamarckista, mediante la instrucción, puede ser *simulada* por la evolución darwinista mediante selección natural.

Este tipo de explicación también tiene lugar en física. Un ejemplo sencillo sería la hipótesis que propuso por primera vez Kant, seguido por Laplace, que trata de explicar el hecho de que todos los planetas de nuestro sistema planetario se mueven en torno al sol en planos más o menos semejantes, no muy alejados y en la misma dirección. Esta "Hipótesis de la nebulosa" (como solía llamarla Spencer) supone como situación inicial típica una nebulosa en rotación a partir de la cual se van formando los planetas por procesos de condesación (o, según Spencer, de diferenciación e integración). De tal modo, la teoría explica o simula lo que a primera vista parecía una ordenación planificada conscientemente. [Podría decirse aquí que la hipótesis de la nebulosa de Kant o Laplace podría perfeccionarse e incluso sustituirse por una hipótesis semejante a la de la "supervivencia". Según esta hipótesis, un sistema de planetas que se mueven en planos muy divergentes o que se mueven en direcciones parcialmente distintas sería menos estable, en diversos órdenes de magnitud, que un sistema como el nuestro; de modo que sólo hay una pequeña probabilidad de encontrar un sistema de un tipo menos estable.] Otro ejemplo físico sería el siguiente: la teoría gravitatoria de Newton opera con fuerza de atracción que actúa a distancia. [G. L. Le Sage publicó en 1782 una teoría que explicaba esta acción a distancia newtoniana, simulándola. En su teoría no hay fuerzas de atracción, sino simples cuerpos que *empujan* a otros cuerpos.] [18] Puede decirse que la teoría gravitatoria de Einstein muestra de qué modo se puede simular el sistema newtoniano mediante un sistema explicativo en el que no hay ni empujes ni fuerzas de atracción. Ahora bien, es importante que la explicación simulada —es decir, la teoría de Newton— pueda ser descrita como una *aproximación* a la teoría de Einstein y a la verdad. La teoría de la selección natural procede de un modo semejante. En cualquier caso particular, parte de una situación simplificada en un modelo —situación formada por determinada especie en unas condiciones ambientales dadas— e intenta mostrar por qué, en esta situación, ciertas mutaciones poseen un valor para la supervivencia. Así, aunque el lamarckismo sea falso, como parece que es, será considerado por los darwinistas como una primera aproximación al darwinismo [18].

La dificultad real del darwinismo es el problema, de sobra conocido, de explicar los desarrollos evolucionistas que *aparentemente* están *dirigidos a un fin,* como nuestros ojos, recurriendo a un número de pasos muy pequeños, pues, según el darwinismo, cada uno de estos pasos es el resultado de una mutación puramente accidental. Es difícil explicar que todas esas

---

[18] [G. L. Le Sage (traducido por Abbot: "The Newtonian Lucretius"), *Annual Report of the Smithsonian Institution,* 1898, págs. 139-60.]

[19] [En la conferencia original, dos de los pasajes de este párrafo aparecían en otra parte (aproximadamente en la página siguiente).]

mutaciones accidentales e independientes hayan tenido un valor de supervivencia. [Es, especialmente, el caso de la conducta heredada de Lorenz.] El "efecto-Baldwin" —es decir, la teoría de un desarrollo puramente darwinista que simula el lamarckismo— me parece un paso importante hacia la explicación de tales desarrollos.

Creo que fue Samuel Butler el primero que vio claramente esta dificultad, condensándola en la pregunta: "¿Casualidad o Astucia?", en el sentido de "¿Accidente o Plan?". De un modo semejante, podemos considerar el sistema bergsoniano de la Evolución Creadora como un comentario a esta dificultad: su *élan vital* no es más que el nombre que él daba a todo lo que causase o controlase aquellos cambios *aparentemente dirigidos a un fin.* Toda explicación animista o vitalista de este tipo es evidentemente *ad hoc* y completamente insatisfactoria. Con todo, se la puede reducir a algo mejor —como hizo Darwin al mostrar que las explicaciones teleológicas se pueden simular— mostrando, por consiguiente, que se trataba de una aproximación a la verdad —o al menos, a otra teoría sostenible—. (Trataré de formulas dicha teoría en la tercera parte de esta conferencia.)

Habría que añadir aquí algunas palabras sobre la forma lógica de la teoría de la selección natural. Es un tema muy interesante que me hubiese gustado exponer aquí con cierta extensión. Pero sólo me es posible señalar uno o dos puntos con brevedad.

La teoría de la selección natural es de carácter *histórico*: construye una situación para mostrar que, dada dicha situación, es muy probable que ocurran aquellas cosas cuya existencia deseamos explicar.

Para decirlo con mayor exactitud, la teoría de Darwin es una explicación histórica *generalizada.* Lo que esto quiere decir es que la situación no es *única,* sino *típica.* Por tanto, algunas veces puede ser posible construir un *modelo* simplificado de la situación.

Tal vez, deba decir aquí con brevedad que lo que tengo por la idea central de Darwin —su intento de explicar los cambios genéticos que conducen a una adaptación mejor en el sentido de *más oportunidades de supervivencia para el animal o planta individuales*— ha sufrido últimamente un eclipse. Se debe en gran medida a la búsqueda de exactitud matemática, que ahora está tan de moda, así como al intento de definir estadísticamente el valor de supervivencia de hecho (de un gene u otra unidad genética de determinada población).

Mas la supervivencia o el éxito, en el sentido de un aumento numérico, se puede deber a dos circunstancias distintas. Una especie, por ejemplo, puede tener éxito o prosperar tras haber logrado aumentar su velocidad, sus dientes, sus habilidades o su inteligencia; o puede prosperar sencillamente por haber conseguido aumentar su fecundidad. Está claro que un aumento suficiente de la fecundidad, que depende básicamente de factores genéticos o de la disminución del período de inmadurez, puede tener el mismo valor de supervivencia, o aún mayor, que, digamos, un aumento de habilidad o inteligencia.

Desde este punto de vista, puede resultar algo difícil comprender por qué la selección natural habría de producir algo más que un aumento general de las tasas de reproducción y la supresión de todas las razas que no fueran las más fértiles [14]. [Puede haber muchos factores distintos implicados en los procesos que determinan las tasas de reproducción y mortalidad, por ejemplo, las condiciones escológicas de la especie, su interrelación con otras especies y el equilibrio entre dos (o más) poblaciones.] Mas, sea como sea, creo que sería posible superar las dificultades considerables que se plantean a la hora de medir *el éxito en la adaptación de los organismos individuales* de una especie substrayendo, tal vez, su valor de fecundidad (su tasa de nacimiento) de su aumento total de población (su tasa de supervivencia). En otras palabras, propongo que digamos que la especie *A está mejor adaptada que la B* (en sentido darwinista y lamarckiano) si, por ejemplo, sus poblaciones aumentan del mismo modo a pesar de que *A* tenga una tasa de nacimiento inferior a la de *B*. En tal caso, podríamos decir que los *miembros individuales* de una especie *A* son, por término medio, más capaces de sobrevivir que los de la especie *B* o que están mejor adaptados a su medio que los de *B*.

Sin una distinción de este tipo (que se puede elaborar estadísticamente) lo más seguro es que perdamos de vista los problemas originales de Lamarck y Darwin, especialmente el poder explicativo de la teoría darwinista —su capacidad de explicar la adaptación y los desarrollos pretendidamente orientados a un fin, mediante la selección natural que *simula* una evolución de corte lamarckiano.

Para finalizar la segunda parte de mi conferencia, he de recordarles que, como ya he indicado antes, no creo en la inducción. Opino que Hume ha mostrado de un modo concluyente la invalidez de la inducción, si bien creía que, aunque inválida e injustificable racionalmente, los animales y los hombres la utilizan de modo generalizado. No creo que sea cierto. Estimo que lo cierto es que procedemos por el método de *seleccionar* anticipaciones, expectativas o teorías —es decir, por el método de ensayo y supresión de errores que se ha interpretado frecuentemente como inducción porque *simula la inducción*—. Creo que el venerable mito de la inducción ha provocado en el pensamiento biológico una buena dosis de dogmatismo. También ha dado lugar a la denuncia de los que son a veces motejados de "científicos de gabinete" —es decir, los teóricos—. Pero nada malo hay en los gabinetes, que han acogido fielmente a Kepler, Newton, Maxwell y Einstein; a Bohr, Pauli, de Broglie, Heisenberg y Dirac; y a Schrödinger, tanto en sus especulaciones físicas como en las biológicas.

---

[14] Esta es una de las innumerables dificultades de la teoría de Darwin a las que parecen ciegos algunos darwinistas. Desde este punto de vista es particularmente difícil de comprender la transición de los organismos unicelúlares a los pluricelulares que se enfrentan a nuevas dificultades peculiares para su reproducción, especialmente para la supervivencia tras la reproducción, y que introducen en la vida algo nuevo: a saber, la muerte, pues todos los individuos multicelulares mueren.

Lo digo con tiento porque yo no soy ni siquiera un biólogo de gabinete, sino algo peor, un simple filósofo de gabinete.

Pero después de todo, también lo era Herbert Spencer, cuyo nombre, he de admitir, estoy explotando descaradamente como tapadera de mis propios desaguisados en el campo de la especulación biológica.

3. Una conjetura: "dualismo genético"

Entro ahora en la tercera parte, la fundamental, de mi conferencia —la presentación de una conjetura o hipótesis que, de resistir la crítica, tal vez pueda reforzar la teoría de la selección natural, aunque se mantiene estrictamente dentro de los límites lógicos del contexto del neodarwinismo (o, si ustedes prefieren, de la "Nueva Síntesis") ortodoxo.

Naturalmente, mi conjetura es una hipótesis histórica generalizada: consiste en construir una situación típica en la que la selección natural pueda producir los resultados que deseamos explicar con su ayuda.

El proplema a resolver es el viejo problema de la ortogénesis versus mutación accidental e independiente —el problema de Samuel Butler de *la casualidad o la astucia*. Surge de la dificultad de comprender de qué modo puede resultar de la cooperación puramente accidental de las mutaciones independientes un órgano complicado como el ojo.

Brevemente, mi solución del problema consiste en la hipótesis según la cual en muchos, si no en todos, los organismos cuya evolución plantea este problema —tal vez haya que incluir algunos organismos de una escala muy baja— podemos distinguir más o menos tajantemente, al menos, *dos partes distintas*: grosso modo, una *parte que controla la conducta*, como el sistema nervioso central, y una parte *ejecutiva*, como los órganos de los sentidos y las piernas, junto con sus estructuras sustentadoras.

Esta es, para decirlo rápidamente, la situación que supone mi conjetura. Se puede combinar con la suposición del neodarwinismo ortodoxo según el cual los cambios por mutación en una de estas partes serán independientes por regla general, aunque tal vez no siempre, de los cambios mutacionales en la otra parte.

Esta hipótesis relativa a dicha situación postula un *dualismo* que se parece muchísimo al dualismo del cuerpo y la mente. Sin embargo, es compatible con la versión más radical del materialismo mecanicista, así como con la forma más radical de animismo, ya que lo único que pide mi hipótesis dualista —que tal vez pueda denominar "*dualismo genético*"— se puede expresar del siguiente modo:

En los casos que queremos explicar, podemos considerar que ciertas disposiciones o tendencias heredadas como las de auto-conservación, búsqueda de alimento, huída del peligro, adquisición de habilidades por imitación, etcétera, están sujetas a mutaciones que, por regla general, no inducen ningún cambio significativo en ningún órgano corporal ni siquiera en los órganos de los sentidos, si exceptuamos (a lo sumo) aquellos órganos que son

los portadores genéticos de las disposiciones o tendencias mencionadas.

Antes de pasar a explicar las consecuencias de esta hipótesis, he de decir inmediatamente que puede que la hipótesis del dualismo genético sea falsa. Sería falsa si los genes (o cualquier otra cosa que cumpla las funciones de estas unidades de control de la herencia) que controlan, por ejemplo, el desarrollo embriológico del ojo humano fuesen siempre los mismos que controlan nuestra curiosidad visual innata —nuestra disposición o tendencia a utilizar los ojos lo más posible en todo tipo de situaciones en que haya la suficiente iluminación para poder ver. Para decirlo de otro modo: si nuestra tendencia innata a *utilizar* los ojos, oídos, manos, piernas, etc., se trasmitiese siempre por herencia exactamente del mismo modo que nuestra *posesión* de ojos, oídos, manos, piernas, etc., entonces mi hipótesis dualista sería falsa. También sería falsa si fuese un error separar tajantemente el *uso* y la *posesión* de un órgano —si, por ejemplo, la posesión y el uso no fuesen más que dos abstracciones distintas de lo que biológica o genéticamente no es más que una y la misma realidad. Denominaré a esta suposición *monismo genético* o *hipótesis monista.*

Creo que la aceptación táctica de algo semejante a este monismo genético es la responsable de que mi hipótesis dualista no se haya (al menos que yo sepa) desarrollado y criticado plenamente hasta ahora. La aceptación de una hipótesis monista tal vez se haya visto favorecida por el hecho de que el problema fundamental de la teoría de la evolución era explicar el origen de las especies —es decir, el origen de la diferenciación de los órganos de los animales y las plantas, más bien que el origen de los tipos específicos de conducta o tendencias comportamentales.

Sea como sea, discutiré ahora el funcionamiento de mi hipótesis dualista sirviéndome de un *modelo mecánico.* Más exactamente, utilizaré un servomecanismo —una máquina— en lugar de un organismo en desarrollo. Pero antes de eso, quiero dejar bien sentado que mi conjetura no es idéntica a dicho modelo y que quienes acepten mi conjetura no se comprometen en absoluto a aceptar el punto de vista según el cual los organismos son máquinas. Además, mi modelo no contiene elementos correspondientes a todos los elementos relevantes de la teoría. Por ejemplo, no incluye ningún mecanismo para producir mutaciones u otros cambios genéticos, por la sencilla razón de que no es ese mi problema.

Tomaré como modelo un avión —por ejemplo, un avión de combate— guiado por un piloto automático. El avión que manejamos aquí está construido para unos fines determinados y el piloto automático está dotado de un cierto número de reacciones incorporadas en su diseño que equivalen a "instrucciones" para atacar a un enemigo más débil, apoyar al compañero en el ataque y la defensa, esquivar a un enemigo más fuerte, etc. Las partes mecánicas del piloto automático de que dependen estas "instrucciones" constituyen la base física de lo que denominaré la *estructura-propositiva* de mi modelo.

Además, el piloto automático lleva incorporada la base física de lo que

denominaré su *estructura de destreza* y que consiste en cosas tales como mecanismos de estabilización, mecanismos para interpretar detectores a fin de identificar y dintinguir amigos de enemigos, controles de dirección, controles de funciones a realizar, etc. No suponemos que la estructura propositiva y la estructura de destreza sean claramente distinguibles. Ambas juntas constituyen lo que propongo denominar *estructura de disposición central* del piloto automático o, si ustedes prefieren, su "mente". El sistema físico —los interruptores, cables, válvulas, baterias, etc., incluyendo los que continen las "instrucciones" del piloto automático— pueden considerarse la base física de su estructura de disposición central o de su "mente". En lo que sigue llamaré sencillamente "piloto automático" a este *sistema físico.*

Sabemos que se puede dotar a este servomecanismo de cierta disposición a "aprender" —por ejemplo, a mejorar algunas de sus habilidades— por ensayo y error. Pero, para empezar, podemos dejar esto de lado. En vez de ello, supondremos inicialmente que la estructura propositiva y la estructura de destreza son rígidas, estando ajustadas a los órganos ejecutivos del avión como, por ejemplo, la potencia del motor.

Supongamos ahora que nuestro avión de combate es reproductible —no importa si se autorreproduce o si se fabrica en una factoría copiando sus diversos elementos físicos— estando sometido a mutaciones accidentales. Agrupemos en cuatro clases las mutaciones posibles.

(1) Mutaciones que afectan al piloto automático.

(2) Mutaciones que afectan a un órgano —digamos, el timón o motor— controlado por el piloto automático.

(3) Mutaciones que afectan a un órgano autorregulador que no está controlado por el piloto automático, por ejemplo, un termostato independiente que regula la temperatura de los motores.

(4) Mutaciones que afectan a la vez a dos o más órganos.

Ahora está claro que en un organismo complicado como éste casi todas las mutaciones accidentales serán desventajosas y, en su mayoría, letales. Por tanto, suponemos que habrán de ser eliminadas por selección natural. Esto mismo se puede decir especialmente de las mutaciones accidentales que afectan a más de un órgano —por ejemplo al piloto automático— *más* otro órgano. Tales mutaciones están abocadas a ser desfavorables; la probabilidad de que ambas sean favorables o incluso complementarias será casi cero.

Esta es una de las mayores diferencias existentes entre mi hipótesis dualista y la monista. Según la hipótesis monista, la mutación favorable a un órgano, por ejemplo un aumento de potencia en uno de los motores, siempre tendrán un *uso* favorable y eso es todo lo que pasa. Toda mutación favorable es improbable, pero no es preciso que su probabilidad sea inmensamente pequeña. Pero de acuerdo con la hipótesis dualista, un cambio favorable en un órgano, en muchos casos sólo será *potencialmente* favorable. Para que represente una diferencia respecto al estado anterior, la mejora ha de ser *utilizada* y este nuevo uso ha de depender de un cambio accidental *comple-*

mentario en la estructura de disposición central. Mas la probabilidad de dos cambios accidentales semejantes, a la vez complementarios e independientes, tiene que ser insignificante.

Por tanto, a primera vista, puede parecer que el modelo dualista sólo puede aumentar las dificultades de una teoría que opere solamente con la selección, lo cual puede que sea otra de las razones por las que la mayoría de los darwinistas parecen haber abrazado tácitamente una hipótesis monista.

Pongamos un ejemplo. Supongamos que una mutación da una mayor potencia a todos los motores, de tal manera que el avión pueda volar más aprisa. Podemos considerar que se trata de algo favorable tanto para atacar al enemigo como para esquivarlo. También podemos suponer que su estructura propositiva inducirá al piloto automático a hacer pleno uso del aumento de potencia y velocidad. Mas su estructura de destreza se ajustará a los viejos motores y a la vieja velocidad y, dado que hemos supuesto que el piloto no puede "aprender" en el sentido de mejorar su destreza, la velocidad será demasiado rápida para él con lo que, de acuerdo con mi hipótesis dualista, el avión se estrellará. El monismo genético, por otra parte, supondrá que con el aumento de la potencia de los motores se produce por sí mismo un aumento de destreza, puesto que no es más que otro aspecto de lo mismo —según la suposición de que, por razones genéticas, no hemos de separar un órgano de su uso.

Recordarán ustedes nuestras cuatro posibilidades de cambio por mutación:

(1) Cambio de la estructura del piloto automático.
(2) Cambio en un órgano directamente controlado por el piloto.
(3) Cambio en un sistema con auto-control.
(4) Cambio en más de un órgano al mismo tiempo.

Como hemos visto, el caso (4) —es decir, un cambio en más de un órgano— es algo que podemos dejar de lado, tanto los partidarios de la hipótesis dualista como los de la monista, ya que los cambios favorables de este tipo son demasiado improbables.

Podemos abordar el caso (3) —es decir, un cambio en un órgano con autocontrol— considerando que tal órgano o es un sistema dualista reducido al que hay que aplicar de nuevo nuestra hipótesis dualista, o bien se le aplica la hipótesis monista y, por tanto, se desarrolla según la teoría al uso.

Lo más probable es que el caso (2) —es decir, un cambio en un órgano directamente controlado por el piloto— sea desfavorable, como muestra nuestro ejemplo del aumento de la potencia de los motores y velocidad, a pesar del hecho de que semejante mutación habría de ser favorable desde el punto de vista de una hipótesis monista.

Nos queda, por tanto, el caso (1), el de los cambios por mutación en la estructura de disposición central heredada. Mi tesis es que los cambios favorables en esta estructura no crean ninguna dificultad especial. Por ejemplo,

una mutación favorable en la estructura propositiva puede inducir al avión a escapar del enemigo en mayor número de casos que antes; o tal vez la disposición opuesta pueda ser la que resulte favorable (es decir, la disposición a atacar al enemigo en un mayor número de casos). No sabemos cuál será más favorable, pero, según nuestros supuestos, la selección natural se encargará de descubrirlo.

Lo mismo ocurre con las habilidades. Sabemos que un piloto humano puede mejorar sus habilidades sin cambiar la estructura de su avión. Esto muestra la posibilidad de introducir mutaciones favorables en la estructura de destreza del piloto automático, sin necesidad de cambios complementarios en el resto de la estructura. Naturalmente, las mutaciones favorables son siempre poco probables, pero sabemos que un piloto humano, sin cambiar su avión, puede adoptar nuevos fines, desarrollar nuevas habilidades y no estrellarse. Algunos de esos nuevos fines y habilidades pueden ser favorables desde el punto de vista de la autoconservación; por tanto, también pueden tener éxito nuevos fines y habilidades del piloto automático.

Llegamos, así pues, al primer resultado siguiente: si partimos de un organismo dualista en el que están perfectamente equilibradas la estructura de disposición central de control y la estructura ejecutiva controlada, entonces parece probable que las mutaciones en la estructura de disposición central sean un poco menos letales que las mutaciones en los órganos ejecutivos controlados (incluso las potencialmente favorables).

Nuestro segundo resultado fundamental es el siguiente. Una vez que un nuevo fin, tendencia o disposición, una nueva habilidad o modo de comportarse ha evolucionado en la estructura de disposición central, este hecho influirá sobre los resultados de la selección natural, de tal modo que las mutaciones que antes eran desfavorables (aunque potencialmente favorables) se hacen efectivamente favorables con tal de que apoyen la tendencia últimamente establecida. *Mas lo que esto quiere decir es que la evolución de los órganos ejecutivos se verá dirigida por aquella tendencia o finalidad y, por tanto, estará "orientada a un fin".*

Podemos ilustrar esto considerando dos tipos de mutaciones favorables de la estructura de disposición central: aquellas con las que se *mejoran* lo que podíamos llamar fines y habilidades amplios y aquellas con las que se *especializan*.

Ejemplo del primer tipo son las mutaciones que introducen fines que son tan sólo indirectamente favorables, como la mutación que introduce un fin, una tendencia, un deseo de mejorar alguna habilidad del organismo. Ahora bien, una vez que se ha establecido una mutación de este tipo, puede tornarse más favorable otra mutación que haga más flexible la estructura de destreza y, mediante tales mutaciones de la estructura de destreza, el organismo podrá adquirir la disposición a "aprender" en el sentido de mejorar su destreza mediante ensayo y error.

Además, una vez lograda una estructura de disposición central más flexible, las mutaciones en los órganos ejecutivos, que de otro modo serían

letales, como puede ser una mayor velocidad, pueden tornarse extremadamente favorables aunque antes fuesen desfavorables.

Lo que aquí interesa es ver que las mutaciones en la estructura central pueden ser *directoras*. Es decir, sólo se conservarán aquellas mutaciones de los órganos ejecutivos que se adapten a las tendencias generales establecidas previamente por los cambios de la estructura central.

Se puede decir algo por el estilo del segundo tipo de cambio; es decir, de los cambios *especializadores* de la estructura central. Los cambios en el medio pueden favorecer la restricción de la estructura propositiva. Por ejemplo, si sólo se puede utilizar fácilmente *un* tipo de alimento —tal vez un tipo no demasiado favorecido en un principio— puede ser altamente favorable un cambio de gusto (es decir, un cambio en la estructura propositiva). Este cambio de finalidad puede llevar a una especialización de todo el organismo, por ejemplo, las habilidades para la consecución de alimento, así como la forma de sus órganos. Para poner un ejemplo, esta teoría sugeriría que el pico y la lengua especializados del pájaro carpintero se han desarrollado por selección, *después* de haber empezado a cambiar sus gustos y hábitos alimenticios y no a la inversa. Además, podemos decir que si el carpintero hubiese desarrollado su pico y su lengua antes de cambiar sus gustos y habilidades, el cambio hubiese sido letal: no hubiese sabido qué hacer con sus órganos.

O, para tomar un ejemplo lamarckiano clásico, la jirafa: según mi teoría, sus tendencias o hábitos tróficos tienen que haber cambiado *antes* que su cuello, de lo contrario un cuello más largo no tendría ningún valor de supervivencia.

Debo dejar ya de exponer mi teoría y decir algunas palabras sobre su poder explicativo.

Para decirlo con brevedad, mi hipótesis dualista nos permite, en principio, aceptar no sólo el lamarckismo simulado, sino también un vitalismo y animismo simulados, con lo que "explica" estas teorías como primeras aproximaciones. Por eso nos permite explicar en principio la evolución de los órganos complejos, como el ojo, mediante muchos pasos que van hacia una dirección definida. Además, como dicen los vitalistas, la dirección puede estar determinada por una tendencia de aspecto mental —por la estructura propositiva o de destreza que puede desarrollar una tendencia o deseo de utilizar el ojo, así como una destreza en la interpretación de los estímulos recibidos a su través.

Al mismo tiempo, no hay razón para pensar que la hipótesis monista tenga que ser siempre falsa. Puede que en el transcurso de la evolución se desarrollen distintos tipos de organismos en mayor o menor medida dualistas o monistas por lo que respecta a su mecanismo genético. De este modo, tal vez podríamos explicar por lo menos algunas irrupciones de cambios evolucionistas aparentemente dirigidos a fines. Por el contrario, otros cambios que están menos orientados a un fin se pueden explicar suponiendo que nos enfrentamos al desarrollo de estructuras genéticamente monistas.

Quizá sea éste el lugar más idóneo para confesar que me vi llevado a

formular esta conjetura acerca del dualismo genético al encontrarme embrollado con un caso que, prima facie, constituye una refutación de mi conjetura y al intentar descubrir por qué me resultaba tan sorprendente. Se trata del caso de la mutación de cuatro alas (*tetraptera*) de la *drosophila*, la famosa mosca de la fruta de dos alas. Lo que me tenía perplejo era lo siguiente: ¿por qué no se estrellaban los mutantes de cuatro alas? ¿Cómo podían tener la habilidad necesaria para utilizar sus cuatro alas? Tal vez este caso no refute efectivamente mi conjetura, pero parece bastante posible que sea así. (Tal vez la estructura alada del insecto es en gran medida autorreguladora o constituye una parte monista de un animal esencialmente dualista, o quizás la mutación sea un atavismo —como se supone que es— con lo que la estructura de destreza —no así la estructura propositiva— relativa al uso de las cuatro alas ha sobrevivido de modo atávico como la más antigua, cambiando probablemente de un modo gradual de cuatro a dos alas.) A parte de la preocupación por este caso, también me impulsaron de un modo fundamental las consideraciones acerca de la evolución del hombre, del lenguaje humano y del árbol del conocimiento humano.

Para terminar, mencionaré un punto que habla muy en favor del dualismo genético: los estudiosos del comportamiento animal han mostrado la existencia de un comportamiento innato complejo —comportamiento que implica la utilización muy hábil, altamente especializada y coordinada de diversos órganos. Me parece difícil, si no imposible, creer que tal comportamiento no sea más que otro aspecto de la estructura anatómica de los diversos órganos que participan en ella.

A pesar de éste y otros argumentos en contra de la hipótesis monista, no creo que mi propia hipótesis dualista se pueda contrastar fácilmente. Sin embargo, no creo que sea incontrastable. No obstante, antes de poder discutir seriamente las posibles contrastaciones, la hipótesis ha de ser examinada críticamente desde el punto de vista de la consistencia, de si resuelve —caso de ser verdadera— los problemas que trata de resolver y de si puede ser mejorada simplificándola y haciéndola más precisa. Por el momento, la ofrezco sencillamente como una posible línea de pensamiento.

SUPLEMENTO

EL PROMETEDOR MONSTRUO COMPORTAMENTAL

La conferencia anterior fue pronunciada hace diez años, en 1961. Algunas de sus ideas —la teoría de las mutaciones comportamentales como puntas de flecha— fueron desarrolladas más ampliamente en "Sobre Nubes y Relojes". Mas, aunque estoy intensamente interesado por la teoría evolucionista, no soy un experto en ninguno de sus campos y un especialista me recomendó que no publicase la Conferencia Spencer.

No obstante, a lo largo de todos estos años, me ha parecido que *distinguir entre las bases genéticas de* (1) *las finalidades o preferencias,* (2) *de las habilidades y* (3) *de las herramientas anatómicas ejecutivas* puede constituir una contribución importante a una teoría evolucionista de corte darwinista. Me ha parecido que aquello que he denominado "dualismo genético" (podría haberlo llamado "pluralismo genético") ofrece una explicación de las tendencias genéticas u "ortogénesis".

Me ha parecido que mejoraba la teoría que Richard B. Goldschmidt había propuesto mediante sus famosos "monstruos prometedores". Creo que es útil comparar ambas teorías.

Goldschmidt (1878-1958) publicó en 1940 un libro, titulado *The Material Basis of Evolution* [1], en el que ponía de relieve que las numerosas variaciones pequeñas de Darwin planteaban inmensas dificultades. En primer lugar, existe una tendencia a volver a una población promedio, aunque hayan aparecido mutaciones. En segundo lugar, en todos los experimentos de selección, nos encontramos con la gran dificultad de corregir cambios más allá de ciertos límites definidos: los intentos de ir más allá llevan casi invariablemente a la esterilidad y la extinción.

Ambos argumentos plantean dificultades a la teoría darwinista que hace partir la evolución de unas pocas formas vivas originales —tal vez de una solamente—. Con todo, es precisamente esta teoría la que queremos explicar, un fenómeno en favor de cuya realidad hay una gran cantidad de elementos de juicio empíricos.

La explicación ortodoxa consiste en decir que, si tomamos períodos de tiempo muy grandes, entonces, es posible acumular variaciones pequeñas y que, sobre todo, la separación geográfica suele impedir el restablecimiento de una población media. Goldschmidt consideraba que estas ideas eran insuficientes y, sin romper con la idea de la selección natural, abandonó la idea de que todo cambio revolucionario hubiese de ser explicado en términos de un número ingente de variaciones muy pequeñas. Supuso que, de vez en cuando, aparecen mutaciones grandes que, aunque en su mayoría sean letales, por lo que se ven eliminadas, algunas veces sobreviven. De este modo, explicaba tanto el carácter genuino de las diferencias como *también* el carácter obvio de parentesco entre las diversas formas vivas. Consideraba esas grandes mutaciones como "*monstruos prometedores*". Su teoría tiene su parte atractiva: los monstruos aparecen de tarde en tarde. Sin embargo, entraña serias dificultades: normalmente, esas mutaciones serán letales (un organismo es algo demasiado cuidadosamente equilibrado como para soportar grandes cambios accidentales y repentinos) y, aún cuando no lo sean, la posibilidad de retorno a la forma original es muy grande.

Siempre he estado muy interesado por las teorías de Goldschmidt [2]. Tanto es así, que he llamado la atención de I. Lakatos sobre sus "monstruos

[1] Richard B. Goldschmidt, *The Material Basis of Evolution*, Yale University Press, New Haven, 1940.

[2] Véase Richard B. Goldschmidt, *op. cit.*, así como su artículo "Some Aspects of Evolution", *Science, 78,* 1933, págs. 539-47.

prometedores" hasta el punto de que alude a ellos en su artículo "Proofs and Refutations" [3].

Mas, hace tan sólo unos días, mientras leía un nuevo libro crítico, *Darwin Retried,* de Norman Macbeth [4], se me ocurrió que ya era hora de resucitar una nueva versión de los "monstruos prometedores" de Goldschmidt.

El propio Goldschmidt pensaba fundamentalmente —si no exclusivamente— en monstruos anatómicos —organismos dotados de diferencias de índole estructural nada despreciables e incluso drásticas respecto a sus padres. Lo que yo sugiero es que debemos partir de *monstruos etológicos* o comportamentales: organismos cuyas diferencias respecto a sus padres consisten primeramente en *conductas* distintas.

Naturalmente, esta conducta tiene su base genética; pero ésta parece permitir una cierta amplitud de respuestas comportamentales que, quizá, sólo dependan del estado fisiológico momentáneo en el que se encuentra el organismo cuando responde a los estímulos de su medio o, tal vez, de una combinación de estímulos poco corriente o de una variación genética en la disposición comportamental. En todos estos casos pueden aparecer, y de hecho aparecen, nuevas conductas monstruosas sin que se pueda observar ninguna novedad anatómica. La novedad puede tener su base material en un cambio restringido a una parte del sistema nervioso, pero dicho cambio puede deberse a una herida o cualquier otro accidente sin que sea necesario que esté genéticamente determinada. Por otro lado, puede deberse perfectamente a una mutación genética genuina relativa a esa parte del sistema genético específicamente responsable de la conducta; dicha mutación no está necesariamente ligada a un cambio anatómico grande. Finalmente, la novedad de la conducta puede deberse a una novedad real en las circunstancias ambientales, en la ecología del organismo.

En todos estos casos, el monstruo comportamental puede separarse radicalmente de sus padres por lo que atañe a su conducta. Sin embargo, no hay ninguna razón inmediata por la que dicha desviación haya de ser letal. Es evidente que la conducta mostruosa puede alterar el equilibrio del organismo, pero no tiene por qué ser así necesariamente, o bien puede alterarlo de un modo no necesariamente letal (como cuando una mosca que se pasea por el papel en que escribo se empapa las patas de tinta y le cuesta un cierto trabajo limpiarlas).

La novedad y mostruosidad (en el sentido de Goldschmidt) de la conducta tiene muchas menos probabilidades de ser letal que la monstruosidad anatómica. Por otro lado, mediante la selección natural, la conducta monstruosa puede tener las mayores consecuencias sobre la supresión de las variaciones anatómicas.

Para tomar el famoso ejemplo del ojo, la nueva *conducta* que utiliza las manchas sensibles de luz (ya existentes) puede aumentar en gran medida

---

[3] I. Lakatos, "Proofs and Refutations", *B. J. P. S., 14,* 1963, pág. 24.

[4] Norman Macbeth, *Darwin Retried,* Gambit Incorporated, Boston, 1971; véase especialmente el capítulo 17.

su valor selectivo que quizás antes fuese despreciable. De este modo, el *interés* por ver puede fijarse con éxito genéticamente, convirtiéndose en el elemento rector de la evolución ortogenética del ojo; hasta las menores mejoras en su anatomía pueden ser valiosas selectivamente si la estructura propositiva y la de destreza las utilizan suficientemente.

Propongo, por tanto, una versión del darwinismo según la cual los monstruos comportamentales desempeñan una función decisiva. Si una novedad comportamental tiene éxito, lleva a la selección de aquellos nichos ecológicos que, a su vez, operarán selectivamente; es decir, operarán sobre esas novedades comportamentales, ejerciendo una presión selectiva en una dirección parcialmente predeterminada: en la dirección determinada por un *fin* indeterminado genéticamente posible; por ejemplo, el gusto por un nuevo tipo de alimento o el placer de utilizar regiones de la piel sensibles a la luz. De este modo, podemos llegar a la ortogénesis que era, después de todo, el problema principal de Goldschmidt.

Incluso propiedades tales como el carácter juguetón juvenil, pueden haberse tornado útiles en un mundo cambiante en el que la mostruosidad comportamental —es decir, variabilidad más posible ortogénesis— puede favorecer la supervivencia.

De este modo, incluso se podría explicar el (frecuente) papel desempeñado por los cambios (genéticos e incluso indeterminados genéticamente) de la estructura propositiva y, secundariamente, de la estructura de destreza sobre los cambios de la estructura anatómica con base genética. En términos generales, la estructura anatómica sólo puede cambiar despacio, razón por la cual sus cambios serán insignificantes si no están guiados por cambios en la estructura propositiva y de destreza. Por tanto, en principio, se podrá explicar en términos darwinistas la evolución de un aparato genético que establezca la primacía de ambas estructuras.

Como se puede ver, esta teoría darwinista de los prometedores monstruos comportamentales no sólo "simula" el lamarckismo, sino también el vitalismo bergsoniano.

8. CONSIDERACION REALISTA DE LA LOGICA, LA FISICA Y LA HISTORIA*

El hombre, dicen algunos filósofos modernos, está alienado de su mundo: es un extraño atemorizado en un mundo que él no ha construido. Tal vez sea así; en todo caso, lo mismo se puede decir de los animales e incluso de las plantas. También ellos nacieron hace mucho en un mundo físico-químico que ellos no habían creado. Mas, aunque no hayan creado su mundo, las cosas vivas lo han cambiado, convirtiéndolo en algo completamente distinto y además han rehecho la pequeña zona del universo que los vio nacer. Tal vez, el mayor de esos cambios haya sido el producido por las plantas, que han transformado radicalmente la composición química de la atmósfera terrestre. El siguiente en magnitud, tal vez venga dado por los logros de algunos animales marinos constructores de arrecifes de coral, así como de islas y cadenas montañosas de caliza. Finalmente, llegó el hombre que durante mucho tiempo no transformó el medio de un modo notable, si exceptuamos su contribución al aumento de las zonas desérticas mediante la tala de árboles. Naturalmente, construyó unas pocas pirámides, pero hasta más o menos el siglo pasado no empezó a hacer la competencia a los corales formadores de arrecifes. Aún más recientemente, ha comenzado a echar a perder la obra de las plantas elevando paulatina, aunque progresivamente, la proporción de dióxido de carbono de la atmósfera.

Por tanto, no hemos hecho nuestro mundo. Hasta el presente, ni siquiera lo hemos cambiado mucho en comparación con los cambios conseguidos por los animales marinos y las plantas. Con todo, hemos creado un nuevo tipo de producto que, pasado el tiempo, promete producir en nuestra parcela del mundo cambios tan grandes como los producidos por nuestros antecesores, las plantas productoras de oxígeno o los corales constructores de islas. Estos nuevos productos, que indudablemente son obra nuestra, son nuestros mitos, nuestras ideas y, especialmente, nuestras teorías científicas acerca del mundo en que vivimos.

Pienso que debemos considerar estos mitos, ideas y teorías como algunos de los productos más característicos de la actividad humana. Como las herramientas, son órganos que evolucionan fuera de nuestra piel: son artefactos exosomáticos. Por tanto, hemos de contar especialmente entre estos productos característicos lo que se suele llamar "conocimiento humano",

---

* Artículo basado en la Disertación Inaugural del Primer Coloquio Internacional celebrado en la Universidad de Denver del 16 al 20 de mayo de 1966. Publicado por primera vez en W. Yourgrau y A. D. Breck (eds.), *Physics, Logic and History*, Plenum Press, 1970, págs. 1-30.

entendiendo "conocimiento" en un sentido objetivo e impersonal según el cual podemos decir que el conocimiento está en los libros, se almacena en las bibliotecas o se enseña en la universidad.

Normalmente, al aludir al conocimiento humano, estaré pensando en este sentido objetivo de la palabra "conocimiento". Esto nos permitirá considerar al conocimiento producido por los hombres como algo análgo a la miel que producen las abejas: las abejas hacen la miel, la almacenan y la consumen. Además, la abeja individual que consume miel, por regla general, no consume únicamente la porción que ella ha producido: también consumen miel los zánganos que no producen absolutamente nada (para no hablar de este tesoro de miel almacenada que las abejas pierden en beneficio de los osos y los apicultores). También es interesante tener en cuenta que las obreras, para conservar la capacidad de producir más miel, han de consumir una parte que normalmente producen otras.

Todo esto se aplica cumplidamente, con ligeras modificaciones, a las plantas productoras de oxígeno así como a los hombres que producen teorías: también nosotros, no sólo producimos, sino que consumimos teorías. Para seguir produciendo, hemos de consumir las teorías de otros y algunas veces, tal vez, las nuestras.

"Consumir" significa aquí, sobre todo, "digerir", como en el caso de las abejas. Pero significa también algo más: consumir teorías, hayámoslas producido nosotros mismos u otras personas, significa también criticarlas, cambiarlas y a veces incluso demolerlas para sustituirlas por otras mejores.

Todas estas operaciones son necesarias para el *desarrollo de nuestro conocimiento;* naturalmente, también aquí entiendo el conocimiento en sentido objetivo.

Parece como si este desarrollo del conocimiento humano, el desarrollo de nuestras teorías, fuese el causante de que nuestra *historia humana* sea un capítulo tan radicalmente nuevo de la historia del universo y de la vida sobre la tierra.

Estas tres historias —la del universo, la de la vida sobre la tierra y la del hombre, junto con el desarrollo de su conocimiento— son ellas mismas, como es natural, otros tantos capítulos de nuestro conocimiento. Por consiguiente, el último de estos capítulos —es decir, la historia del conocimiento— estará formado por conocimiento acerca del conocimiento. Habrá de contener, al menos de un modo implícito, teorías acerca de teorías; sobre todo, teorías acerca del modo en que se desarrollan las teorías.

Por tanto, antes de seguir desarrollando el tema, presentaré un esquema tetrádico general que ha mostrado reiteradamente su utilidad para describir el desarrollo de las teorías. Es como sigue:

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$$

Aquí "*P*" significa "problema", "*TT*", "teoría tentativa" y "*EE*", "(intento de) eliminar errores", sobre todo mediante la discusión crítica. Mi esquema tetrádico trata de mostrar que el resultado de la crítica o supre-

sión de errores aplicado a una teoría es, por regla general, el surgimiento de un nuevo problema o de varios de ellos. Una vez que los problemas han sido resueltos y una vez que sus soluciones han sido examinadas correctamente, tienden a engendrar hijos problemáticos: nuevos problemas, normalmente más profundos e incluso de mayor fecundidad que los viejos. Esto se puede ver especialmente en la física y creo que la mejor manera que tenemos de calibrar el progreso realizado en una ciencia consiste en medir la distancia que hay entre P_1 y P_2 por lo que respecta a su profundidad y esperabilidad: las mejores teorías tentativas (y todas lo son) son las que plantean los problemas más profundos e inesperados.

Mi esquema tetrádico se puede elaborar de diversas maneras; por ejemplo:

$$P_1 \begin{cases} \nearrow & TT_a \rightarrow EE_a \rightarrow P_{2a} \\ \rightarrow & TT_b \rightarrow EE_b \rightarrow P_{2b} \\ \searrow & TT_n \rightarrow EE_n \rightarrow P_{2n}. \end{cases}$$

De esta forma, el esquema indica que, si podemos, hemos de proponer varias teorías que intenten resolver un problema dado y que hemos de examinar críticamente cada una de nuestras soluciones tentativas. Luego, vemos que cada una de ellas da lugar a nuevos problemas, por lo que hemos de seguir aquellas que prometan problemas más novedosos e interesantes: si el nuevo problema, digamos P_{2b}, resulta no ser más que el viejo P_1 disfrazado, entonces decimos que nuestra teoría lo único que consigue es *desplazar* un tanto *el problema;* incluso, en algunos casos, eso constituirá una objección decisiva contra la teoría tentativa TT_b.

Esto muestra que la eliminación de errores es sólo una *parte* de la discusión crítica: nuestra discusión crítica de teorías tentativas rivales puede ponerlas en comparación y valorarlas desde muchos puntos de vista distintos. Pero, *como es natural,* el punto decisivo es siempre el mismo: ¿En qué medida nuestra teoría resuelve bien sus problemas, es decir, P_1?

En cualquier caso, una de las cosas que deseamos conseguir es aprender algo nuevo. Según nuestro esquema, una de las cosas que exigimos a una buena teoría tentativa es el carácter progresivo que se descubre con la discusión crítica: la teoría es progresiva si la discusión muestra que *realmente da lugar a ciertas diferencias respecto al problema que queríamos resolver;* es decir, si los últimos problemas que han surgido son distintos de los viejos.

Si los últimos problemas surgidos son distintos, podemos esperar aprender una gran cantidad de cosas cuando procedamos a resolverlos a su vez.

Por tanto, mi sistema tetrádico se puede utilizar para describir la emergencia de nuevos problemas y, en consecuencia, la de nuevas soluciones; es decir, nuevas teorías. Incluso quiero presentarlo como un intento de conferir sentido a la idea eminentemente vaga de emergencia, como un intento de hablar de emergencia de un modo racional. Quisiera decir que se puede aplicar no sólo a la emergencia de nuevos problemas científicos y, consi-

guientemente, a nuevas teorías, sino también a la emergencia de nuevas formas de comportamiento e incluso de nuevas formas de organismos vivos.

Pondré un ejemplo. P_1 puede ser, digamos, un problema determinado relativo a la supervivencia de las especies, como es el problema de la reproducción, de la producción de la prole. Para Darwin, este problema de supervivencia encuentra una buena solución cuando la especie sobrevive; cualquier otra solución tentativa será eliminada por la desaparición tanto de la solución como de la especie.

Según mi esquema, el intento de supresión de errores —es decir, la lucha por la supervivencia— pondrá de manifiesto la debilidad inherente a cada una de las soluciones propuestas en forma de un *nuevo problema.* Por ejemplo, el nuevo problema puede ser el de que los organismos paternos y los de la prole amenacen con asfixiarse mutuamente. Este nuevo problema, a su vez, puede resolverse; por ejemplo, los organismos pueden desarrollar un método para disgregar o dispersar su prole; o el nuevo problema puede resolverse mediante el establecimiento de una economía común que abarque diversos organismos. Tal vez la transición de los organismos unicelulares a los pluricelulares haya procedido de este modo.

Sea como sea, mi esquema muestra que puede haber otras alternativas además de la Darwin —"*sobrevivir o perecer*"— inherentes al proceso de supresión de errores: esta supresión puede dar lugar a nuevos problemas emergentes específicamente relacionados con el viejo *y* con la solución tentativa.

En adelante utilizaré mi esquema, aunque a veces lo haré de un modo implícito; además, aludiré a la emergencia, suponiendo que mi esquema hace lo suficientemente respetable a esta idea para lo que considero una discusión racional. Propongo que consideremos algunos aspectos del desarrollo del conocimiento desde cuatro puntos de vista:

1. Realismo y Pluralismo: Reducción versus Emergencia.
2. Pluralismo y Emergencia en Historia.
3. Realismo y Subjetivismo en Física.
4. Realismo en Lógica.

1. Realismo y pluralismo: reducción versus emergencia

El hombre no sólo produce teorías científicas, sino que también produce muchas otras ideas —por ejemplo, mitos poéticos o religiosos y argumentos de narraciones.

¿Cuál es la diferencia característica que media entre una teoría científica y una obra de ficción? Sostengo que no es que la teoría sea posiblemente verdadera mientras que las descripciones de la narración no lo sean, si bien la verdad y falsedad tienen aquí algo que ver. Creo que la diferencia estriba en que la teoría y la narración están insertas en distintas tradi-

ciones críticas. Han de ser juzgadas mediante normas tradicionales muy diversas (aún cuando dichas normas puedan tener algo en común).

Lo característico de la teoría es que se presenta como una solución a un problema científico; es decir, sea a un problema que ha surgido antes en la discusión crítica de teorías tentativas primitivas o sea (tal vez) a un problema descubierto por el autor de la teoría ahora ofrecida en el ámbito de los problemas y soluciones pertenecientes a la tradición científica.

Sin embargo, no basta decir eso, ya que la tradición científica se caracteriza, o se caracterizaba hasta hace poco, por lo que se podría denominar *realismo científico;* es decir, se inspiraba en el ideal de encontrar *soluciones verdaderas* para sus problemas: soluciones correspondientes a los hechos.

Este ideal regulador, consistente en dar con teorías que se corresponden con los hechos, es lo que convierte la tradición científica en una tradición realista: discrimina entre el mundo de nuestras teorías y el mundo de los hechos al que pertenecen estas teorías.

Además, las ciencias naturales, así como también algunas ciencias sociales, especialmente la historia y la economía, con sus métodos críticos de resolución de problemas han representado durante mucho tiempo nuestros mejores esfuerzos en la resolución de problemas y en el descubrimiento de hechos (por descubrimiento de hechos entiendo, naturalmente, el descubrimiento de enunciados o teorías correspondientes a los hechos). Por tanto, dichas ciencias contienen los enunciados y teorías mejores desde el punto de vista de la verdad; es decir, los que suministran las mejores descripciones del mundo de los hechos o de lo que llamamos "realidad".

Consideremos ahora determinadas relaciones vigentes entre algunas de estas ciencias.

Tomemos, por ejemplo, la física y la química, ciencias que hacen aserciones sobre todas las cosas y estados físicos, incluyendo los organismos vivos.

La física y la química no son muy distintas y no parece que haya mucha diferencia entre los tipos de cosas a que se aplican, excepto en que la química, tal como se entiende normalmente, se torna inaplicable a temperaturas muy elevadas y, quizás también, muy bajas. Por tanto, no sería muy sorprendente el que las esperanzas tan antiguas de reducir la química a la física resultasen ser verdaderas, como parece ocurrir.

Aquí nos encontramos con un verdadero caso paradigmático de "*reducción*". Por *reducción* entiendo, naturalmente, que todos los descubrimientos de la química pueden explicarse plenamente mediante (es decir, ser deducidos de) los principios de la física.

Aunque tal reducción no fuese muy sorprendente, constituiría un gran éxito científico. No sólo sería un ejercicio de unificación, sino también un avance real en la comprensión del mundo.

Supongamos que esta reducción ha sido plenamente realizada. Este hecho nos daría algunas esperanzas de que algún día podamos reducir también toda la biología a la física.

Ahora bien, esto constituiría un éxito espectacular mucho mayor que la reducción de la física a la química. ¿Por qué? Porque el tipo de cosas a que se aplican la física y la química son realmente muy semejantes desde un principio. Piénsese tan sólo en lo difícil que sería decir si la teoría atómica es física o química. De hecho, fue ambas cosas a la vez durante mucho tiempo, siendo este lazo común el que suministra una unidad que puede llevar, o ha llevado tal vez, a su unificación.

Con los organismos vivos la situación es distinta. Sin duda están sometidos a todo tipo de leyes físicas y biológicas. Con todo, prima facie, parece haber alguna diferencia entre organismos vivos y cosas inanimadas. Hemos de admitir que la ciencia nos muestra la existencia de estadios de transición intermedios, así como de sistemas intermedios. Todo esto nos permite esperar que algún día se llevará a cabo la reducción. Por otro lado, no parece en absoluto que las recientes teorías tentativas acerca del origen de la vida sobre la tierra puedan ser contrastadas con éxito y seamos capaces de crear artificialmente organismos vivos primitivos.

Pero ni siquiera esto significaría necesariamente una reducción completa, como muestra el hecho de que los químicos eran capaces de crear todo tipo de sustancias químicas, inorgánicas y orgánicas, antes de comprender cuál era su composición química, por no hablar de su estructura física. Así pues, ni siquiera el control de los procesos químicos por medios puramente físicos equivale a una reducción de la química a la física. Reducir significa mucho más; significa comprensión *teórica*: penetración teórica del nuevo campo en el viejo.

Por tanto, podríamos descubrir una receta para crear algunas formas primitivas de vida a partir de materia inanimada sin comprender teóricamente lo que estamos haciendo. Es evidente que eso constituiría, con todo derecho, un tremendo estímulo para quienes buscan una reducción, pero el camino hasta la reducción podría ser todavía largo y ni siquiera sabemos si es o no transitable: puede no haber una reducción teórica de la biología a la física, del mismo modo que no parece haber una reducción teórica ni de la mecánica a la electrodinámica ni a la inversa.

Si la situación es tal que, por un lado, los organismos vivos pueden originarse a partir de sistemas inanimados por un proceso natural y, por otro, no hay posibilidad de comprensión teórica completa de la vida en términos físicos, entonces tendríamos que considerar la vida como una propiedad *emergente* de los cuerpos físicos o de la materia.

Ahora bien, quiero dejar bien claro que, como racionalista, deseo y espero comprender el mundo, por lo que deseo y espero una reducción. A la vez, considero muy probable que no pueda haber tal reducción; es perfectamente concebible que la vida sea una propiedad *emergente* de los cuerpos físicos.

Lo que aquí quiero decir es que quienes creen en la reducción —quienes por una u otra razón filosófica adoptan *a priori* la posición inicial dogmática según la cual la reducción ha de ser posible— destruyen de algún

modo su propio triunfo, en el caso de que consigan siempre llevarla a cabo. La razón estriba en que tal logro habrá de conseguirse siempre; por tanto, es el suyo un éxito carente de interés, consistente en que los acontecimientos se encarguen de demostrar que estaban en lo cierto.

Sólo quienes afirmen que el problema no puede zanjarse *a priori* podrán considerar que una reducción con éxito constituye un descubrimiento tremendo.

Me he detenido tanto en este punto porque tiene cierta relevancia para la posición que forma el siguiente peldaño de la escala, la emergencia de la conciencia.

Hay filósofos denominados "conductistas radicales" o "fisicalistas" qce creen poseer razones válidas *a priori*, del tipo de la *navaja de Ockham*, para afirmar que las introspecciones de nuestros estados o sucesos mentales no son sino introspecciones e informes acerca de nosotros mismos *qua* sistemas físicos: son informes sobre estados físicos de dichos sistemas.

Dos filósofos que han hablado aquí esta mañana han defendido dicha posición con argumentos brillantes. Se trata de Herbert Feigl y Willard Van Orman Quine. Quisiera hacer algunas consideraciones críticas acerca de sus opiniones.

Quine dice, aludiendo a Carnap y Feigl, que si se puede "lograr" un progreso teórico "postulando estados mentales distintos... tras el comportamiento físico, es bien cierto que... se podría hacer otro tanto postulando en su lugar... ciertos estados y acontecimientos fisiológicos correlativos. Difícilmente puede esgrimirse la falta de una explicación fisiológica detallada de estos estados en contra de su reconocimiento como tales estados del cuerpo humano... Sea como sea los estados corporales existen; ¿por qué añadir los otros?" [1].

Permítaseme señalar que Quine habla aquí como realista: "Sea como sea, los estados corporales existen", nos dice. Sin embargo, desde el punto de vista adoptado aquí, no es lo que yo llamaría un "realista científico": no espera a ver si la ciencia consigue llevar a cabo aquí una reducción, como quizá ocurra algún día; en vez de ello, aplica la navaja de Ockham [2], señalando que las *entidades* mentales no son necesarias para la teoría.

¿Mas quién sabe lo que quiere decir aquí Ockham, o quién sea, con la palabra necesidad? Si las entidades mentales o, mejor dicho, los estados mentales existieran —yo, por mi parte, no dudo de que existan— entonces postular estados mentales es necesario para explicarlos verdaderamente; si además resultan algún día reducidos a estados físicos, eso será un éxito tremendo. Pero no habrá ningún éxito si rechazamos su existencia, señalando simplemente que podemos explicar las cosas sin ellos por el mero expediente de limitarnos a las cosas físicas y a su comportamiento.

---

[1] W. V. Quine, *Word and Object*, 1960, pág. 264 [traducción castellana de Manuel Sacristán Luzón, *Palabra y Objeto*, Barcelona, Labor, 1968, pág. 272].

[2] W. V. Quine, *From a Logical Point of View*, segunda edición revisada, 1961, página 2 [traducción castellana de Manuel Sacristán Luzón, *Desde un Punto de Vista Lógico*, Barcelona, Ariel, 1962, pág. 26].

Para resumir mi argumento en pocas palabras, diré que las especulaciones filosóficas de índole materialista o fisicalista son muy interesantes e incluso pueden indicar el camino hacia una fecunda reducción científica. Pero serán teorías francamente tentativas (como pienso que son las teorías de Feigl). No obstante, algunos fisicalistas no consideran que sus teorías sean provisionales, sino que las toman como propuestas de expresarlo todo en un lenguaje fisicalista; además, creen que estas propuestas tienen mucho en su favor porque son, sin duda, *convenientes*: algunos problemas inconvenientes, como el problema del cuerpo y la mente, desaparecerán de la manera más conveniente. Así, dichos fisicalistas creen que no se puede dudar de que estos problemas han de ser eliminados como pseudoproblemas.

A esto respondería que por el mismo procedimiento tendríamos que haber eliminado *a priori* todos los estados químicos, así como los problemas relacionados con ellos: podríamos haber dicho que eran obviamente físicos y que no era necesario especificarlos en detalle: lo único que necesitábamos hacer era postular la existencia de algún estado físico correlativo a cada estado químico.

Me parece que está claro que la adopción general de tal propuesta habría provocado una actitud consistente en no buscar la reducción detallada de la química a la física. Sin duda, habría disuelto el problema análogo al del cuerpo y la mente, el problema de la relación entre física y química; pero la solucción habría sido lingüística y por consiguiente, no habríamos aprendido nada acerca del mundo real.

Todo esto me inclina a decir que el realismo debería ser pluralista, al menos provisionalmente y que los realistas deberían suscribir el siguiente postulado pluralista:

Hemos de cuidarnos de resolver o disolver los problemas fácticos lingüísticamente; es decir, por el simple expediente de negarse a hablar de ellos. Por el contrario, hemos de ser pluralistas, al menos de entrada: en primer lugar, deberíamos de hacer hincapié en las dificultades, aunque parezcan insolubles, como a muchos puede parecerles que ocurre con el problema del cuerpo y la mente.

Por tanto, si mediante la reducción científica podemos reducir o eliminar algunas entidades, hagámoslo por todos nuestros medios y felicitémonos por la mayor comprensión así alcanzada.

Así pues, yo diría: *sea como sea, antes de intentar la reducción,* elucidemos con detalle en cada caso concreto los argumentos que hay en favor de la emergencia.

Resumiendo y precisando las consideraciones expuestas en esta sección:

. La reducción de la química a la física, que hoy día parece bien encaminada, puede tomarse como un caso paradigmático de genuina reducción científica que satisface todos los requisitos de una buena explicación científica.

Una reducción "buena" o "científica" es un proceso en el que aprendemos muchas cosas de gran importancia: aprendemos a comprender y ex-

plicar las teorías relativas al campo que se trata de reducir (la química en este caso) y aprendemos también una gran cantidad de cosas relativas a la potencia de las teorías reductoras (la física en este caso).

Aunque no sea seguro, se puede concebir perfectamente que la reducción de la química a la física tenga un éxito completo. También se puede concebir, aunque sea menos probable, que demos algún día con *buenas reducciones* de la biología, incluyendo la fisiología, a la física, así como de la psicología a la fisiología y, por tanto, a la física.

Llamo *reducciones malas* o *ad hoc* al método consistente en construir reducciones valiéndonos de argucias meramente lingüísticas, como ocurre con el método fisicalista que nos sugiere la postulación *ad hoc* de la existencia de estados fisiológicos para explicar una conducta que ya antes habíamos explicado postulando (si bien no de un modo *ad hoc*) estados mentales. O, dicho en otras palabras, mediante la estratagema lingüística consistente en decir que cuando afirmo que ahora creo comprender la ecuación de Schrödinger, lo que hago es informar acerca de un estado *fisiológico* mío.

Este segundo tipo de reducción o utilización de la navaja de Ockham es malo porque nos impide ver el problema. Dicho con la pintoresca y sugestiva terminología de Imre Lakatos, constituye un caso desastroso de "*desplazamiento degenerador del problema*" que impide una buena reducción, el estudio de la emergencia o ambas a la vez.

Para evitar este método desastroso, hemos de tratar, en todos los casos, de aprender lo más posible acerca del campo que deseamos reducir. Puede ser que dicho campo se resista a ser reducido y, en algunos casos, incluso es posible formular argumentos que muestren por qué dicho campo no puede ser reducido. En tal caso, tendremos un ejemplo de emergencia genuina.

Quizá pueda terminar mis comentarios sobre el desplazamiento degenerador del problema que tiene lugar con el conductismo (especialmente el conductismo lingüístico) haciendo la siguiente consideración:

Los conductistas y los materialistas son anti-idealistas que se oponen, con todo derecho, a la fórmula de Berkeley "*esse = percipi*" o

$$ser = ser\ observable$$

Según ellos, "ser" es "ser material", "comportarse como un cuerpo en el espacio y el tiempo". No obstante, puede decirse que inconscientemente se adhieren a la ecuación de Berkely, aunque formulada verbalmente de un modo ligeramente distinto:

$$ser = ser\ observado$$

o, tal vez,

$$ser = ser\ percibido,$$

ya que para ellos sólo existen las cosas que pueden ser observadas. No se dan cuenta de que *toda observación entraña interpretación a la luz de teorías* y que lo que ellos llaman "observable" es lo observable a la luz de

teorías primitivas de vieja factura. Si bien soy partidario del sentido común, también soy partidario de ampliar sus dominios atendiendo a las enseñanzas de la ciencia. En todo caso, *no es la ciencia, sino una dudosa filosofía (o una ciencia periclitada), quien conduce al idealismo, al fenomenalismo y al positivismo; o al materialismo y al conductismo* o a cualquier otra forma de anti-pluralismo.

2. Pluralismo y emergencia en la historia

No hablaré de la historia del universo, sino que me limitaré a decir unas pocas cosas acerca de la historia de la vida sobre la tierra.

Parece que últimamente se han dado los primeros pasos con visos de éxito hacia la reconstrucción de las condiciones bajo las cuales la vida ha *emergido* sobre la tierra y creo que quizá pronto podamos presenciar grandes éxitos en este terreno. Pero, mientras que soy un partidario acérrimo de la emergencia, incluso de la experimental, soy muy escéptico acerca de la reducción, lo cual se debe a algunas de mis ideas acerca de la evolución de la vida.

Me parece que los procesos evolucionistas o los cambios evolucionistas más importantes son tan impredictibles como los procesos históricos o los cambios históricos más señalados. Mantengo esta postura porque siento una fuerte inclinación hacia una visión indeterminista del mundo más radical que la de Heisenberg: mi indeterminismo incluye la tesis de que incluso la física clásica es indeterminista y, por tanto, muy semejante a la de Charles Sanders Peirce o a la de Alfred Landé. Creo que, en gran medida, la evolución avanza de un modo probabilístico bajo condiciones o situaciones problemáticas que cambian constantemente y creo que toda solución tentativa, tenga mayor o menor éxito o carezca totalmente de él, crea una nueva situación problemática. Creo que esto impide una reducción completa así como una comprensión completa de los procesos vitales, si bien no impide un avance constante y fecundo hacia tal comprensión. (Este argumento no habría que tomarlo en el sentido en que Bohr aplica su idea de la complementariedad a los organismos vivos, argumento que me parece muy débil.)

Mas de lo que quiero hablar fundamentalmente en esta sección es de la historia humana, de los avatares de la humanidad. Ya he indicado que en gran medida esto no es más que la historia de nuestro conocimiento, nuestras teorías acerca del mundo y, naturalmente, de las repercusiones de esos productos hechos por nosotros sobre nosotros mismos y nuestras producciones ulteriores.

Es obvio que se puede adoptar una actitud fisicalista o materialista hacia éstos productos teóricos nuestros. Se podría sospechar además que el hincapié que hago en el sentido objetivo del conocimiento —el hincapié que hago sobre las teorías en cuanto que están contenidas en libros reunidos en bibliotecas y que se enseñan en las universidades— pone al descubierto mi simpatía por la interpretación materialista o fisicalista de las teorías. Me re-

fiero a una interpretación que considera que el lenguaje consta de objetos físicos —ruidos o letras impresas— y que piensa que estamos condicionados a responder a dichos ruidos o letras con determinados tipos característicos de conducta física.

Pero nada más lejos de mi ánimo que fomentar semejantes reducciones *ad hoc*. He de admitir que si me viese obligado a tener que elegir entre una concepción subjetivista o personalista del conocimiento humano y una concepción materialista o fisicalista, como la que acabo de bosquejar, elegiría ésta última; pero *no* es en absoluto ésta la alternativa.

La historia de las ideas nos enseña con toda claridad que éstas emergen en contextos lógicos o, si ustedes prefieren, dialécticos [3]. Mis diversos esquemas del tipo

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$$

pueden considerarse racionalizaciones o desarrollos del esquema dialéctico hegeliano: son racionalizaciones porque operan plenamente en el seno del organon clásico de la crítica racional, que se basa en la llamada ley de contradicción; es decir, en la exigencia de eliminar contradicciones allí donde aparezcan. A nivel científico, la supresión de errores procede mediante una búsqueda consciente de contradicciones.

Por tanto, la historia —especialmente la historia de las ideas— nos enseña que para comprender la historia hemos de comprender las ideas y sus relaciones lógicas (o dialécticas) objetivas.

No creo que quien haya penetrado seriamente en cualquier capítulo de la historia de las ideas pueda pensar que vaya a tener éxito la reducción de dichas ideas. Sin embargo, me impondré aquí como meta argüir, no tanto en contra de la posibilidad de tal reducción, cuanto en favor del reconocimiento de las entidades emergentes, así como de la necesidad de reconocer y describir estos *entia* emergentes antes de poder pensar seriamente en su posible eliminación mediante la reducción.

Ya he expuesto en otro lugar [4] uno de los argumentos principales que tengo en favor del carácter emergente de las teorías. Mi argumento depende de la conjetura según la cual hay un aumento genuino del conocimiento científico; o, para decirlo en términos prácticos, que mañana o dentro de un año podemos proponer o contrastar teorías importantes en las que nadie había pensado seriamente hasta entonces. Si hay un aumento del conocimiento en este sentido, entonces no puede ser predictible por medios científicos, pues quien pudiese predecir hoy, por medios científicos, nuestros descubrimientos de mañana, podría realizarlos hoy, lo cual significaría poner fin al desarrollo del conocimiento.

---

[3] K. R. Popper, "¿Qué es la Dialéctica?", en *Conjectures and Refutations*, 1963 [traducción castellana de Néstor Míguez, *El Desarrollo del Conocimiento Científico. Conjeturas y Refutaciones*. Buenos Aires, Paidos, 1967, capítulo 15].

[4] K. R. Popper, *The Poverty of Historicism*, 1957, Prefacio. [Traducción castellana de Pedro Schwartz, *La Miseria del Historicismo*. Madrid, Taurus, 1961. Reeditada en Alianza Editorial. Madrid, 1973.]

Por otra parte, la impredictibilidad en principio siempre se ha tenido como el aspecto más notable de la emergencia, por lo que creo que mi argumento muestra que, en cualquier caso, el desarrollo del conocimiento ha de ser en principio impredictible.

Pero hay otros argumentos en favor del carácter emergente de las teorías o del conocimiento en sentido objetivo. Sólo mencionaré uno de los dos argumentos que hay en contra de la opinión, tan popular como ingenua, según la cual las teorías pueden reducirse a los estados mentales de quienes las producen o comprenden. (No discutiremos si es posible además que estos estados mentales mismos puedan ser reducidos a su vez a estados físicos.)

La idea de que una teoría, en sentido lógico u objeivo, pueda ser reducida a los estados mentales de sus partidarios, por regla general se formula diciendo que la teoría no *es* más que un pensamiento. Mas esto constituye un error trivial asentado en la incapacidad de distinguir dos sentidos de la palabra "pensamiento". En su sentido subjetivo, la palabra "pensamiento" describe una experiencia o proceso mental. Mas dos procesos o experiencias mentales, aunque puedan mantener relaciones causales entre sí, no pueden mantener relaciones lógicas.

Por tanto, si digo que algunas ideas determinadas de Buda coinciden con otras de Schopenhauer o que contradicen ciertas ideas de Nietzsche, entonces no hablo acerca de procesos mentales de pensamiento de esas personas ni de sus interrelaciones. Por el contrario, si digo que Nietzsche estaba influido por ciertas ideas de Schopenhauer, entonces doy a entender que ciertos procesos de pensamiento de Nietzsche estaban influenciados causalmente por su lectura de Schopenhauer. Así pues, de hecho, tenemos dos mundos distintos, el de los *procesos de pensamiento* y el de los *productos* de dichos procesos. Mientras que en el primero pueden tener lugar relaciones *causales,* en el segundo tienen lugar relaciones *lógicas.*

El que ciertas teorías sean incompatibles es un hecho lógico que tiene lugar al margen de que alguien constate o comprenda dicha incompatibilidad. Estas relaciones lógicas puramente objetivas son características de esas entidades que he denominado teorías o conocimiento en sentido objetivo.

Se desprende esto mismo del hecho de que, a menudo, quien produce una teoría no la comprende. Así podría argüirse, sin caer en una paradoja, que Erwin Schrödinger no comprendió plenamente la ecuación de Schrödinger, al menos hasta que Max Born le dio su interpretación estadística, o que la ley de las áreas de Kepler no fue propiamente comprendida por el propio Kepler a quien, al parecer, disgustaba.

De hecho, comprender una teoría es algo similar a una tarea infinita, de manera que podemos decir perfectamente que una teoría nunca se comprende plenamente, aunque haya algunas personas capaces de comprender alguna teoría extremadamente bien. Además, comprender una teoría tiene mucho en común con la comprensión de una personalidad humana. Podemos conocer o comprender muy bien el sistema de disposiciones de una persona; es decir, podemos ser capaces de predecir de qué forma se comportaría en un

cierto número de situaciones distintas. Pero, puesto que hay un número infinito de situaciones posibles de infinita diversidad, la comprensión plena de las disposiciones de una persona no parece posible. Con las teorías pasa algo similar: comprender plenamente una teoría significaría comprender todas sus consecuencias lógicas. Pero éstas son infinitas en un sentido que no es trivial: hay infinitas situaciones de infinita variedad a las que debería aplicarse la teoría; es decir, sobre las que puede tener alcance alguna de sus consecuencias lógicas. Pero muchas de esas consecuencias nunca han sido tomadas en consideración, su posibilidad puede no haber sido descubierta aún. Ahora bien, esto significa que nadie, ni su creador ni cualquiera que intente aprenderla, puede poseer una plena comprensión de todas las posibilidades inherentes a la teoría, lo cual muestra de nuevo que la teoría, en sentido lógico, es algo objetivo que existe objetivamente —un objeto que podemos estudiar, algo que podemos tratar de aprehender—. Decir que las teorías o ideas son productos nuestros a pesar de que no las comprendemos plenamente no es más paradójico que decir que nuestros hijos son un producto nuestro aunque no los comprendemos plenamente o que la miel es un producto de las abejas que ninguna de ellas comprende plenamente.

Por tanto, el estudio de la historia de nuestras teorías o ideas —y podríamos romper una lanza en favor de la opinión según la cual toda la historia humana es en gran medida la historia de nuestras teorías o ideas— nos convertiría a todos nosotros en pluralistas. La razón de ello es que, para el historiador, lo que existen son personas en situaciones problemáticas físicas, sociales, mentales e ideológicas; personas que producen ideas con las que tratan de resolver esos problemas, ideas que tratan de captar, criticar y desarrollar.

El estudioso de la historia de las ideas se encontrará con que éstas poseen una especie de vida (naturalmente, se trata de una metáfora); con que pueden ser mal comprendidas, rechazadas y olvidadas, con que pueden reafirmarse a sí mismas y resucitar. No obstante, podemos decir de un modo no metafórico que no son idénticas a ningún pensamiento o creencia humanos y que pueden existir aunque sean mal comprendidas y rechazadas por todos.

Tal vez todo esto no sea más que una reminiscencia platónica o hegeliana, pero existen diferencias importantes. Las "ideas" platónicas eran conceptos o nociones eternas y fijas; las de Hegel eran conceptos o nociones que se autotransformaban dialécticamente. Las ideas para mí más importantes no son en absoluto conceptos o nociones; no corresponden a palabras, sino a enunciados o proposiciones.

Frente a Platón y Hegel, considero que los ciudadanos más importantes del mundo de las ideas son las *teorías tentativas* acerca del mundo —es decir, las hipótesis junto con sus consecuencias lógicas—. No creo (como Platón) que su extraño carácter atemporal las haga eternas y, por tanto, *más reales* que las cosas que han sido engrendadas y están sujetas a cambio y degeneración. Por el contrario, una cosa susceptible de cambio y aniquila-

ción, precisamente por esto, ha de ser aceptada como real prima facie. Incluso una ilusión es real *qua* ilusión.

Esto es importante con respecto al problema del tiempo y del cambio. Creo que un historiador no puede aceptar la doctrina de que el tiempo y el cambio son ilusiones, doctrina sostenida por algunos grandes físicos y filósofos como Parménides, Weyl y Schrödinger. Nada es más real que un acontecimiento, un suceso, y todo acontecimiento implica algún cambio.

Me parece, no ya improbable, sino imposible que el universo pluralista en que vive el historiador (con sus hombres individuales viviendo sus vidas individuales, intentando resolver sus problemas, engendrando hijos e ideas acerca de ellos, deseando, temiendo y engañándose a sí mismos y a los demás, aunque teorizando siempre y buscando a menudo, no sólo la felicidad, sino también la verdad) pueda "reducirse" plenamente con éxito a uno u otro tipo de monismo. Pero no me ocuparé aquí de ésto. Lo que quiero señalar es que la navaja de Ockham sólo la podemos empezar a emplear seriamente tras haber reconocido la pluralidad de lo que hay en este mundo. Invirtiendo la bella fórmula de Quine [5], sólo vale la pena utilizar la navaja de Ockham en caso de que la barba de Platón sea suficientemente coriacea y esté liada con demasiadas entidades. Es de esperar que el filo de la navaja se melle al usarla para esta tarea correosa. No cabe duda de que la tarea será penosa, pero esto es algo que siempre ocurre.

3. Realismo y subjetivismo en física

Hay dos campos importantes en la física moderna en los cuales los físicos no sólo han permitido la entrada del subjetivismo, sino que además le han otorgado un papel esencial: la teoría de Boltzmann acerca de la subjetividad de la dirección del tiempo y la interpretación de Heisenberg de las fórmulas de indeterminación que fijan un límite inferior al efecto de la interferencia del observador sobre el objeto observado.

Hubo también otra intrusión del sujeto, o del observador, cuando Einstein introdujo al observador en un cierto número de experimentos mentales imaginarios con la intención de ilustrar la relatividad; pero fue el propio Einstein quien expulsó al observador, lenta aunque firmemente, de este terreno.

No discutiré más este punto ni tampoco la teoría subjetiva del tiempo que, mientras trata de decirnos que el tiempo y el cambio son ilusiones humanas, olvida que son ilusiones muy reales que no han sido en absoluto reducidas a otra cosa (y que, según supongo, no son susceptibles de reducción). No discutiré todas estas cosas, puesto que ya lo he hecho hace poco. Sólo deseo decir un par de palabras sobre las fórmulas de Heisenberg y su interpretación.

---

[5] W. V. Quine, *From a Logical Poin of View*, 2.ª edición revisada, 1961, pág. 2. [Véase la traducción mencionada en la nota 2].

Normalmente, estas fórmulas se derivan de un modo bastante complicado; por ejemplo, hay una derivación interesante de Weyl [6] y otra un tanto complicada de Born [7].

Con todo, de hecho, la fórmula de la energía de Heisenberg no depende ni de la mecánica ondulatoria ni de la mecánica matricial de Heisenberg; tampoco se precisan las relaciones de conmutación (que, según Hill [8], no bastan para derivar las fórmulas). Sencillamente, no depende de la nueva mecánica cuántica revolucionaria de 1925-6, sino que se sigue directamente del viejo postulado cuántico de Planck de 1900.

(1) $$E = h\nu.$$

De donde obtenemos inmediatamente

(2) $$\Delta E = h\,\Delta\nu.$$

Mediante el principio de la potencia resolutoria armónica, tenemos que

(3) $$\Delta\nu \approx 1/\Delta t,$$

a partir de (2) y (3) obtenemos

(4) $$\Delta E \approx h/\Delta t,$$

que lleva inmediatamente a

(5) $$\Delta E \Delta t \approx h;$$

es decir, una versión de las llamadas *fórmulas de indeterminación* de Heisenberg.

Del principio de Duane (cuya analogía con el principio de Planck ha puesto de relieve recientemente Alfred Landé*) obtenemos, del mismo modo, la fórmula de Heisenberg para la posición y el momento. Se puede escribir de este modo:

(6) $$\Delta p_i \approx h/\Delta q_i.$$

Según Landé, esto se puede interpretar del siguiente modo: un cuerpo (una retícula o un cristal) dotado de una periodicidad espacial Δq_i puede cambiar su momento p_i según múltiplos de $\Delta p_i \sim h/\Delta q_i$.

De (6) se sigue inmediatamente

(7) $$\Delta p_i\, \Delta q_i \approx h,$$

que resulta ser otra versión de las fórmulas de indeterminación de Heisenberg.

Teniendo en cuenta que la teoría de Planck es estadística, las fórmulas de Heisenberg pueden interpretarse del modo más natural como *relaciones de*

---

[6] H. Weyl, *The Theory of Groups and Quantum Mechanics*, 1931, págs. 72 y 392.
[7] M. Born, *The Natural Philosophy of Cause and Chance*, 1949, págs. 189-91.
[8] E. L. Hill, en *Mind, Matter, and Method, Essays in Philosophy and Science in Honor of Herbert Feigl* (eds. P. Feyerabend y G. Maxwell), 1966, pág. 442.
* Cf. A. Landé, *Nuevos Fundamentos de la Mecánica Cuántica*, trad. V. Sánchez de Zavala, Madrid, Tecnos, 1968. [*N. T.*].

dispersión estadísticas, tal como he propuesto hace más de treinta años [9, 10]. Es decir, no dicen nada acerca de la precisión posible de las mediciones ni sobre los límites de nuestro conocimiento. Mas, si son relaciones de dispersión, nos dicen algo acerca de los límites de homogeneidad de los estados físicos cuánticos y por ende, si bien de modo indirecto, acerca de la predictibilidad.

Por ejemplo, la fórmula $\Delta p_i \Delta q_i \approx h$ (que puede obtenerse a partir del principio de Duane del mismo modo que $\Delta E \Delta t \approx h$ se puede obtener a partir del principio de Planck) lo único que nos dice es que si determinamos la coordinada x de un sistema (un electrón, por ejemplo) entonces, al repetir el experimento, el momento mostrará una dispersión.

Ahora bien, ¿cómo puede contrastarse tal afirmación? Realizando una larga serie de experimentos con un obturador que abra Δx y midiendo, en cada caso particular, el momento p_x. Si estos momentos muestran una dispersión como la que ha sido predicha, entonces la fórmula ha superado la contrastación. Mas lo que esto muestra es que para contrastar las relaciones de dispersión, hemos medido de hecho p_x en todos los casos con una precisión muy superior a Δp_x ya que de otro modo no podríamos hablar de Δp_x como dispersión de p_x.

Experimentos como el descrito se realizan todos los días en los laboratorios de física. Sin embargo, refutan la interpretación indeterminista de Heisenberg, ya que las medidas (aunque no las predicciones basadas en ellas) son más precisas de lo que permite dicha interpretación.

El propio Heisenberg señaló la posibilidad de tales medidas, aunque decía que era "una cuestión de creencia personal" o "gusto personal" el que les otorgásemos o no algún significado y, desde entonces, han sido universalmente desestimadas por carentes de significado. Sin embargo, tienen significado, pues desempeñan una función bien definida: son contrastacionese de las fórmulas en cuestión; es decir, de las fórmulas de indeterminación *qua* relaciones de dispersión.

Por tanto, no hay razón alguna para aceptar la interpretación subjetivista de Heisenberg o la de Bohr de la mecánica cuántica. La mecánica cuántica es una teoría estadística porque son estadísticos los problemas que intenta resolver —intensidades espectrales, por ejemplo—. Por tanto, no hace falta hacer aquí ningún alegato filosófico en pro de su carácter no-causal.

No obstante, puede establecerse la irreductibilidad de las teorías estadísticas a las deterministas (más bien que la incompatibilidad entre ambas). Landé y otros muchos, así como yo mismo, hemos argumentado en este sentido.

Resumiendo, no hay ninguna razón para dudar del carácter realista y

---

[9] K. R. Popper, *The Logic of Scientific Discovery*, 1959, 1968, 1972 (primera edición alemana, 1934). [Traducción castellana de Víctor Sánchez de Zavala, *La Lógica de la Investigación Científica*, Madrid, Tecnos, 1962; véase el capítulo IX.]

[10] K. R. Popper, "Quantum Mechanics without 'The Observer' ", en *Quantum Mechanics and Reality* (ed. M. Bunge), 1967.

objetivo de toda la física. El papel que en la física moderna desempeña el sujeto que observa no es en absoluto distinto del que desempeñaba en la dinámica de Newton o en la teoría de Maxwell del campo eléctrico: el observador es esencialmente el que contrasta la teoría. Para ello, precisa un montón de teorías distintas, teorías rivales y auxiliares, lo cual muestra que no somos tanto observadores cuanto pensadores.

4. Realismo en lógica

No soy partidario de considerar la lógica como una especie de juego. Conozco los llamados sistemas lógicos alternativos y yo mismo he inventado uno, mas los sistemas lógicos alternativos se pueden discutir desde puntos de vista muy distintos. Se podría pensar que es cuestión de preferencia o convención qué lógica se adopta. Estoy en desacuerdo con esta opinión.

Mi teoría, en pocas palabras, es la siguiente. Considero que la lógica es la teoría de la deducción o derivabilidad o como queramos llamarlo. La derivabilidad o deducción entraña esencialmente *la transmisión de la verdad y la retransmisión de la falsedad*: en una inferencia válida la verdad se transmite de las premisas a la conclusión. Esto se puede utilizar especialmente en lo que denominamos "demostraciones". Pero también se retransmite la falsedad de la conclusión a (por lo menos) una de las premisas, lo cual se emplea en las infirmaciones o refutaciones, esepecialmente en *discusiones críticas*.

Dadas unas premisas y una conclusión, si mostramos que la conclusión es falsa, suponiendo que la inferencia sea válida, sabemos que una al menos de nuestras premisas ha de ser falsa. De este modo se usa constantemente la lógica en la discusión crítica, ya que en una discusión crítica tratamos de mostrar que hay algo que no está en regla en una afirmación. Tratamos de mostrarlo, aunque podemos no tener éxito: la crítica puede ser contestada válidamente mediante contracrítica.

Lo que quisiera sostener es (1) que la crítica es el instrumento metodológico más importante y (2) que si se responde a la crítica diciendo "No me gusta su lógica: su lógica puede estar muy bien para usted, pero yo prefiero otra distinta según la cual esta crítica no es válida", entonces se mina el método de la discusión crítica.

Ahora, yo distinguiría entre dos usos básicos de la lógica; a saber, (1) su uso en las ciencias demostrativas —es decir, las ciencias matemáticas— y (2) su uso en las ciencias empíricas.

En las ciencias demostrativas la lógica se utiliza fundamentalmente para las demostraciones —para la transmisión de verdad— mientras que en las ciencias empíricas se usa casi exclusivamente de modo crítico —para la retransmisión de la falsedad—. Naturalmente, también entran aquí las matemáticas aplicadas en las que utilizamos implícitamente las demostraciones de la matemática pura, mas el papel de las matemáticas en las cien-

cias empíricas es un tanto ambiguo en varios aspectos. (Hay un maravilloso artículo de Schwartz en este sentido [11].)

Así, en las ciencias empíricas, la lógica se utiliza fundamentalmente para la crítica; es decir, para la refutación. (Recuérdese mi esquema $P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$.)

Lo que quiero decir ahora es lo siguiente. Si queremos utilizar la lógica en un contexto crítico, deberíamos utilizar una lógica muy fuerte, la lógica más fuerte, por así decir, que tengamos a mano, ya que queremos que nuestra crítica sea *rigurosa*. Para que la crítica sea rigurosa hemos de emplear todo el aparato; hemos de emplear todas las escopetas que tengamos, todas las balas son importantes. No importa que seamos hipercríticos: si lo somos, se nos responderá con una contracrítica.

Así pues (en las ciencias empíricas), deberíamos utilizar toda la lógica clásica o bivalente. Si no la utilizamos y nos retiramos al uso de una lógica más débil —digamos, la lógica intuicionista o una lógica trivalente (como la que sugería Reichenbach en relación con la teoría cuántica)— entonces, afirmo, no somos lo bastante críticos; es señal de que algo está podrido en el estado de Dinamarca (que en, este caso, es la teoría cuántica en su interpretación de Copenhague, como ya he indicado).

Por el contrario, consideremos ahora las demostraciones. Todo matemático sabe lo interesante que es demostrar un teorema sirviéndose de un *aparato mínimo*. Una demostración que emplee medios más fuertes de lo necesario es matemáticamente insatisfactoria, siendo siempre objeto de interés encontrar las suposiciones más débiles o los medios mínimos posibles para la demostración. En otras palabras, queremos que la demostración sea no sólo suficiente —es decir, válida—, sino también, si es posible, necesaria, en el sentido de recurrir al mínimo de suposiciones. Admito que esta postura es un tanto sofisticada. En unas matemáticas no sofisticadas nos sentimos dichosos y agradecidos si podemos demostrar algo, mas en unas matemáticas más sofisticadas lo que realmente nos interesa es saber qué es *necesario* para demostrar un teorema.

Así, pues, si podemos demostrar teoremas matemáticos con métodos más débiles que la batería completa de la lógica clásica, entonces logramos algo extraordinariamente interesante desde un punto de vista matemático. Por tanto, en la teoría de la demostración nos interesa debilitar lo más posible la lógica clásica, pudiendo, por ejemplo, introducir la lógica intuicionista u otra más débil, como la lógica positiva, para averiguar hasta dónde podemos llegar sin emplear toda la batería.

Creo, dicho sea de paso, que la expresión "lógica intuicionista" es inadecuada. No es sino el nombre de una versión interesante y un tanto debilitada de la lógica clásica, inventada por Brouwer y formalizada por Heyting.

---

[11] J. Schwartz, "The Pernicious Influence of Mathematics on Science", en *Logic, Methodology, and Philosophy of Science* (eds. E. Nagel, P. Suppes y A. Tarski), 1962, páginas 356-60.

No es mi intención hablar aquí en favor de la teoría filosófica denominada intuicionismo, aunque querría romper una lanza en pro de la lógica de Brouwer-Heyting. Confío, no obstante, en que no se piense que estoy tratando de defender en ningún sentido la autoridad de la intuición en filosofía, lógica o donde sea. Dejando de lado, por el momento, la lógica de Brouwer, podría decirse que el intuicionismo es la doctrina según la cual las intuiciones no sólo son importantes, sino, en general, *dignas de confianza.* Contra ello pienso que, si bien las intuiciones son muy importantes, por regla general no resisten la crítica. Por tanto, no soy intuicionista. No obstante, la lógica de Brouwer o la llamada "lógica intuicionista" es muy importante desde el punto de vista de esta discusión, pues no es más que una parte genuina, y por ende una versión debilitada, de la lógica clásica; es decir, toda inferencia válida desde el punto de vista de la lógica intuicionista, lo es también desde el de la lógica clásica, si bien no ocurre a la inversa: hay inferencias que se pueden obtener de manera válida en la lógica clásica, aunque no sean válidas en lógica intuicionista. Por tanto, si puedo demostrar un teorema (sólo demostrado hasta ahora por medios clásicos) con la lógica intuicionista, he hecho un descubrimiento matemático real; pues los descubrimientos matemáticos no consisten exclusivamente en dar con nuevas demostraciones de nuevos teoremas, sino que también consisten en dar con nuevas demostraciones de viejos teoremas. Además, la nueva demostración de un teorema será especialmente interesante si emplea medios más débiles que la vieja demostración. *A fortiori,* siempre podremos construir una demostración de lo mismo utilizando medios más fuertes; con todo, el descubrimiento de una demostración más débil es un logro matemático real.

Así pues, la lógica intuicionista constituye un enfoque muy interesante de las matemáticas porque intenta demostrar el mayor número de teoremas matemáticos posibles valiéndose de medios lógicos reducidos.

La lógica intuicionista posee además otra ventaja: se puede mostrar que en ella no es demostrable la llamada "ley del tercio excluso" (si bien es una fórmula bien formada del sistema). También se puede mostrar que si en un sistema cualquiera no es demostrable una fórmula bien formada, entonces el sistema ha de ser consistente. Hablando en general, cuanto más débiles son los medios lógicos empleados, menor es el peligro de inconsistencia —el peligro de derivar una contradicción—. Pero la lógica intuicionista puede considerarse como un intento de atestiguar que nuestros argumentos son consistentes y no incurrimos en inconsistencias ocultas, paradojas o antinomias. No quiero entrar ahora en el problema de lo segura que es dicha lógica debilitada; pero, como es obvio, es al menos un poco más segura que la lógica clásica completa. No doy por supuesto que sea siempre segura, pero no me ocuparé de esta cuestión. Lo que ahora quiero decir es lo siguiente: si se quiere demostrar o establecer algo, hay que emplear medios débiles, si bien para derrocarlo —es decir, para criticarlo— hay que utilizar medios fuertes. Evidentemente podría decirse, "Mire usted, puedo refutarlo incluso con medios débiles; ni siquiera necesito utilizar toda la lógica intuicionista".

No importa, lo fundamental es que el racionalista da la bienvenida a *cualquier* crítica, si bien puede responder criticando la crítica.

Ahora bien, este enfoque racionalista es un enfoque realista de la lógica. En primer lugar, porque considera la lógica en conexión parcial con la metodología de las ciencias naturales que, como he intentado mostrar, es una cuestión realista. En segundo lugar, y esto es muy importante, porque considera que la inferencia lógica transmite la verdad o retransmite la falsedad, es decir, se ocupa de la idea de verdad.

Yo diría que unos de los logros más importantes de Alfred Tarski consiste en haber introducido en la lógica dos ideas que la han convertido en un asunto verdaderamente mucho más realista. La primera de ellas es la idea de Tarski (en parte anticipada por Bolzano) según la cual la consecuencia lógica es la transmisión de la verdad. Yo diría que la segunda es la rehabilitación de la teoría de la verdad como correspondencia, la rehabilitación de la idea de que la verdad no es más que la correspondencia con los hechos.

Creo que en este punto discrepo un tanto de Quine, pues pienso que esta idea de Tarski debería interpretarse como una destrucción del relativismo y considero correcta la pretensión de Tarski de que su teoría es una teoría "absolutista" de la verdad. Para explicar este extremo contaré una vez más una historia muy vieja dándole un giro ligeramente nuevo. La vieja historia es la de las tres teorías básicas de la verdad. El nuevo giro es la supresión en esta historia de la palabra "verdad" y, con ella, de la apariencia de estar tratando acerca de palabras o definiciones verbales. Sin embrgo, para suprimirla hace falta una discusión preparatoria.

De las tres teorías fundamentales de la verdad, la más antigua era la de la correspondencia, la teoría de que la verdad es la correspondencia con los hechos o, para decirlo de un modo más exacto, que un enunciado es verdadero si (y sólo si) corresponde a los hechos o los describe adecuadamente. Creo que es ésta la teoría rehabilitada por Tarski. La segunda, es la llamada teoría de la coherencia: consideramos que un anunciado es verdadero si (y sólo si) es coherente con el resto de nuestro conocimiento. La tercera teoría dice que la verdad es la utilidad pragmática.

Ahora bien, la teoría de la coherencia se presenta en multitud de versiones de las que sólo mencionaré dos. Según la primera, la verdad es coherencia con nuestras creencias o, más exactamente, un enunciado determinado es verdadero si es concorde con el resto de nuestras crencias. Encuentro esto un tanto desconcertante, pues no deseo incluir las creencias en la lógica por razones de sobra conocidas. (Si Pedro cree p y si p y q son interdeductibles, podríamos decir que Pedro está lógicamente obligado a creer q: Sin embargo, puede que él no sepa que p y q son interdeductibles, con lo que de hecho puede no creer q.)

Según la segunda versión de la teoría de la coherencia, hemos de aceptar como verdadero un enunciado determinado, que no sabemos si es o no verdadero, si (y sólo si) concuerda con los enunciados que hemos aceptado anteriormente. Esta versión tiene como consecuencia hacer que nuestro co-

nocimiento sea manifiestamente conservador: el conocimiento "atrincherado" difícilmente puede ser derribado.

La teoría de la utilidad pragmática se ocupa especialmente del problema de las teorías en las ciencias naturales, como la física. Afirma que deberíamos aceptar como verdadera una teoría física si en las contrastaciones y otras aplicaciones resulta ser pragmáticamente útil o fructífera.

Me propongo echar mano de una especie de treta que consiste en lo siguiente. Ahora en seguida dejaré de hablar de la *verdad* hasta casi el final del artículo. Ya no preguntaré ¿"Qué es la verdad?". Hay varias razones para ello. La principal es que creo que deberían suprimirse las preguntas del tipo "¿Qué es?" o, "¿Qué son?" o, en otras palabras, todas las preguntas verbales o definitorias. Considero que las preguntas "¿Qué es?" o "¿Qué son?" constituyen pseudo-preguntas; no todas lo parecen, pero creo que todas lo son. Creo que no se deberían plantear preguntas tales como "¿Qué es la vida?", "¿Qué es la materia?" o "¿Qué es la mente?". Son típicas preguntas estériles.

Por tanto, creo que también deberíamos descartar la pregunta: "¿Qué es la verdad?"

Podríamos denominar "antiesencialismo" a la primera de las razones (que acabo de mencionar) para descartar la pregunta "¿Qué es la verdad?" Mi segunda razón es aún más importante: también deberíamos evitar, como si de una plaga se tratase, el discutir acerca del significado de las palabras. Discutir el significado de las palabras es uno de los juegos favoritos de la filosofía pasada y presente: los filósofos parecen haberse aficionado a la idea de que las palabras y su significado son algo importante, constituyendo una ocupación filosófica especial.

Para conveniencia de ustedes presentaré aquí, una vez más —en la p - gina siguiente— una tabla que ya he utilizado antes. (Véase la pág. 121 más arriba.)

A la izquierda, tenemos *palabras o conceptos y sus significados* y a la derecha, *enunciados, proposiciones o teorías y su verdad.*

Ahora bien, la experiencia de toda una vida en este terreno me ha enseñado que siempre se debería intentar salir del lado izquierdo de la tabla para mantenerse en el derecho. Siempre deberíamos mantenernos en el terreno de las afirmaciones o teorías, así como en el problema de su verdad. Nunca deberíamos dejarnos envolver en cuestiones verbales o de significado, nunca deberíamos interesarnos por las palabras. Si se nos desafía con la pregunta de si una de las palabras que empleamos significa esto o tal vez lo otro, deberíamos responder; "No lo sé, no me interesan los significados, si usted quiere, aceptaré con gusto *su* terminología". Esto no causa ningún mal. Nunca deberíamos disputar acerca de las palabras ni vernos envueltos en cuestiones de terminología. Deberíamos mantenernos siempre alejados de la discusión de conceptos. Lo que realmente nos interesa, nuestros problemas reales, son problemas fácticos o, en otras palabras, problemas acerca de teorías y su verdad. Nos interesan las teorías y de qué modo soportan

la discusión crítica y nuestra discusión crítica está controlada por nuestro interés en la verdad.

LAS IDEAS

es decir, las

DESIGNACIONES, TÉRMINOS *o* CONCEPTOS		PROPOSICIONES, ENUNCIADOS *o* TEORÍAS
	pueden formularse con	
PALABRAS		AFIRMACIONES
	que pueden ser	
SIGNIFICATIVAS		VERDADERAS
	y su	
SIGNIFICADO		VERDAD
	se puede reducir a	
CONCEPTOS INDEFINIDOS		PROPOSICIONES PRIMITIVAS
	mediante	
DEFINICIONES		DERIVACIONES

El intento de establecer (más bien que reducir) por estos medios su SIGNIFICADO / VERDAD *lleva a una regresión infinita*

Tras decir esto, trataré de dejar de usar la plabra "verdad". Ya no nos planteamos las preguntas: "¿Es la verdad una correspondencia?", "¿Es coherencia?", "¿Es utilidad?". Siendo así, ¿cómo formular nuestro problema real?

Dicho problema sólo puede ser drásticamente planteado señalando el hecho de que todos los que se oponen a las teorías de la correspondencia hacen una *afirmación*. Todos afirman que no puede existir esa correspondencia entre un enunciado y un hecho. Esta es su afirmación central. Dicen que esta idea carece de significado (o que es indefinible, cosa que en mi opinión no importa, pues no importan las definiciones). En otras palabras, todo el problema surge por culpa de las dudas o del escepticismo acerca de la correspondencia: si hay o no algo así como correspondencia entre un enunciado y un hecho. No cabe duda de la seriedad de estas dudas (sobre todo teniendo en cuenta la paradoja del mentiroso).

También es evidente que si no fuese por estas dudas, quienes sostienen la teoría de la coherencia o la de la utilidad pragmática no tendrían nada que objetar. Nadie niega la importancia de la utilidad pragmática de cosas tales como el poder predictivo. Mas si hubiese algo así como *correspondencia de una teoría con los hechos*, entonces es evidente que eso sería más importante que la simple auto-consistencia y también mucho más impor-

tante que la coherencia con el "conocimiento" (o "creencia") previo; pues, si una teoría corresponde a los hechos y contradice algún conocimiento previo, entonces dicho conocimiento ha de ser descartado.

De un modo similar, si existe algo así como la correspondencia de una teoría con los hechos, es evidente que por regla general la teoría que corresponda a los hechos será muy útil; más útil, *qua* teoría, que otra que no corresponda a los hechos. (Por otro lado, puede ser muy útil a un criminal que comparece ante un jurado aferrarse a una teoría que no corresponda a los hechos; mas no es *éste* el tipo de utilidad a que se refieren los pragmatistas; sus puntos de vista plantean un problema muy embarazoso para ellos: me refiero a la pregunta: "¿Util para qué?".)

Aunque soy enemigo del pragmatismo como filosofía de la ciencia, admito de buen grado que el pragmatismo ha hecho hincapié en algo muy importante: el problema de si una teoría tiene alguna aplicación, si tiene, por ejemplo, poder predictivo. Como he dicho en algún lugar, la *praxis* es inestimable para el teórico a la vez como espuela y como brida: es una espuela porque nos sugiere nuevos problemas y una brida porque nos lleva a tierra y a la realidad cuando nos perdemos en imaginarios vuelos teóricos superabstractos. Aunque haya que admitir todo esto, es evidente que la posición pragmatista se verá superada por otra realista si podemos decir con pleno sentido que un enunciado o una teoría puede o no corresponder a los hechos.

Por tanto, la teoría de la correspondencia no niega la importancia de las teorías de la coherencia y de la utilidad pragmática, aunque implica que no son lo suficientemente buenas. Por otro lado, esas teorías afirman la imposibilidad o carencia de significado de la teoría de la correspondencia.

Así pues, sin mencionar siquiera la palabra "verdad" ni preguntar "¿Qué significa la verdad?", podemos ver que el problema central de toda esta discusión no es el problema verbal de definir "verdad", sino el siguiente problema substancial: ¿puede existir un enunciado o teoría que corresponda a los hechos o que no corresponda a ellos?

Tras las dudas relativas a la posibilidad de hablar acerca de la correspondencia se ocultan varios argumentos de peso.

Ante todo, esta idea de correspondencia plantea paradojas o antinomias. En segundo lugar, están los incontables intentos fallidos de decir con mayor exactitud en qué consiste la correspondencia entre un enunciado y un hecho. Hay un intento de Schlick para quien la correspondencia ha de explicarse mediante una relación de uno a uno entre el enunciado lingüístico y el hecho; es decir, mediante la unicidad. Afirma que un enunciado es "verdadero" o corresponde a los hechos si mantiene con los hechos del mundo una relación de uno a uno o única: la no-correspondencia o "falsedad" equivale a ambigüedad. Evidentemente, este punto de vista es inaceptable, puesto que varios enunciados vagos y ambiguos (como "en algún lugar de América hay poca gente") pueden corresponder a los hechos; y viceversa, toda proposición general o teoría que corresponda a los hechos, corresponde a muchos hechos, de modo que no hay una relación biunívoca.

Además, un enunciado que no corresponda a los hechos puede no ser ambiguo. Un asesino puede decir de manera nada ambigua: "Yo no he matado". No hay ninguna ambigüedad en esta afirmación, aunque no corresponda a los hechos. Es evidente que el intento de Schlick de explicar la correspondencia hierra el tiro. Otro intento aún peor es el de Wittgenstein [12]. Wittgenstein sugería que una proposición es una pintura de la realidad y que la correspondencia es una relación muy similar a la que hay entre el surco de un disco y los sonidos que señala: una especie de relación proyectiva entre hechos y enunciados. Es fácil mostrar que estas opiniones son insostenibles. Podemos recordar la famosa anécdota de Livingstone cuando en una ocasión mantuvo una entrevista con un rey negro asistido por un intérprete. Al preguntarle "¿Cómo está usted?" el rey negro respondió con una palabra y el intérprete empezó a hablar y hablar y hablar, durante diez minutos, traduciéndole Livingston aquella palabra mediante una larga historia sobre las penas del rey. A continuación, Livingstone preguntó si el rey precisaba asistencia médica, a lo que el rey contestó hablando, hablando y hablando, cosa que el intérprete tradujo con una sola palabra: "No".

Sin duda esta anécdota es inventada, pero está bien traída e ilustra la debilidad de la teoría proyectiva del lenguaje, especialmente si se la toma como correspondencia entre un enunciado y un hecho.

Pero no termina aquí la cosa. La cuestión es aún más seria. Efectivamente, tras haber formulado su teoría, Wittgenstein dijo que era imposible discutir las relaciones entre el lenguaje y la realidad o discutir el lenguaje en absoluto. (Puesto que el lenguaje no se puede discutir con el lenguaje.) Es un campo en el que las palabras nos fallan. Su expresión favorita para indicar el fracaso de las palabras es "se muestra por sí mismo". Según esto, todo intento de profundizar más en las relaciones entre lenguaje y realidad o de discutir el lenguaje de un modo más profundo se expone a no tener significado. Aunque en el Prólogo comienza diciendo "la *verdad* de los pensamientos aquí comunicados me parece intocable y definitiva", termina afirmando: "quien me comprende acaba por reconocer que [las proposiciones del *Tractatus*] carecen de sentido". * (Porque carece de significado hablar acerca del lenguaje.) No cabe duda de que esto alude especialmente, entre otras cosas, a su teoría de la proyección. Así, la consideración de que sus lectores constatarán la falta de significado de lo que él dice, confirma lo que los oponentes de la teoría de la correspondencia han dicho siempre de ella, a saber, que carece de significado hablar acerca de la correspondencia entre un enunciado y un hecho.

Volvemos así al problema real: ¿hay o no una teoría aceptable de la correspondencia? ¿Podemos o no hablar significativamente de la correspondencia entre un enunciado y un hecho?

---

[12] L. Wittgenstein, *Tractatus Logico-Philosophicus*, 1922. [Traducción castellana de E. Tierno Galván, Madrid, Revista de Occidente, 1956; reeditado en Alianza Editorial, Madrid, 1973.]

* Trad. citada, Prólogo (págs. 32-3) y Proposición 6.54 (págs. 202-3). [*N. T.*]

Lo que yo digo es que Tarski ha rehabilitado la teoría de la correspondencia. Creo que se trata de un importante logro de gran alcance filosófico. Lo digo, porque muchos filósofos (por ejemplo, Max Black) han negado que el resultado de Tarski tenga importancia filosófica.

La clave para la rehabilitación de la teoría de la correspondencia viene dada por una observación muy simple y obvia que hizo el propio Tarski. A saber, si deseo hablar acerca de la correspondencia entre un enunciado, *E*, y un hecho, *H*, tengo que recurrir a un lenguaje en el que pueda hablar acerca de ambos: enunciados del tipo de *E* y hechos del tipo de *H*. Aunque esto parece terriblemente trivial, resulta ser decisivo. Significa que el lenguaje en el que hablamos para explicar la correspondencia ha de tener los medios necesarios para *referirse* a enunciados y *describir* hechos. Si disponemos de un lenguaje dotado de ambos medios, de modo que pueda referirse a enunciados y describir hechos, entonces en dicho lenguaje —el *metalenguaje*— podremos hablar sobre la correspondencia entre enunciados y hechos sin ninguna dificultad, como vamos a ver.

Un metalenguaje es un lenguaje en el que podemos hablar acerca de otro lenguaje. Por ejemplo, una gramática alemana escrita en castellano utiliza este idioma como metalenguaje para hablar sobre el alemán. El lenguaje sobre el que hablamos valiéndonos del *metalenguaje* (en este caso el castellano) se suele denominar "el *lenguaje objeto*" (en este caso el alemán). Lo típico del metalenguaje es que contiene *nombres* (metalingüísticos) de las palabras y enunciados del lenguaje objeto, así como *predicados* (metalingüísticos), como "nombre (del lenguaje objeto)", "verbo (del lenguaje objeto)" o "enunciado (del lenguaje objeto)". Para que un metalenguaje sea suficiente para nuestros propósitos ha de contener también, como señala Tarski, los métodos usuales necesarios para hablar, por lo menos, acerca de los *hechos* de que puede hablar el lenguaje objeto.

Esto es lo que ocurre cuando utilizamos el castellano como metalenguaje para hablar acerca del alemán (como lenguaje objeto sobre el que investigamos).

Por ejemplo, hemos de poder decir en el metalenguaje castellano cosas tales como:

Las palabras alemanas "*Das Gras ist grün*" constituyen un enunciado en alemán.

Por otro lado, hemos de poder describir en nuestro metalenguaje (castellano) el hecho que describe el enunciado alemán "*Das Gras ist grün*". Describimos ese hecho en castellano diciendo, sencillamente, que la hierba es verde.

Ahora podemos formular en el metalenguaje un enunciado sobre la *correspondencia* con *los hechos de un enunciado del lenguaje objeto* del modo siguiente. Podemos hacer la afirmación: *El enunciado alemán "Das Gras ist grün" corresponde a los hechos si, y sólo si, la hierba es verde.* (o: "*...sólo si de hecho la hierba es verde*").

Aunque sea algo trivial, es importante constatar lo siguiente: en nuestra

afirmación, las palabras "*Das Gras ist grün*", *entrecomilladas*, funcionan como un nombre metalingüístico (es decir, *castellano*) del enunciado *alemán;* por otro lado, las palabras castellanas "la hierba es verde" aparecen en nuestra anterior afirmación *sin* ningún tipo de comillas: no funcionan como nombre de un enunciado, sino simplemente como descripción de un *hecho* (o de un hecho supuesto).

Esto posibilita que nuestra afirmación exprese una relación entre un *enunciado* (alemán) y un *hecho.* (El *hecho* no es ni alemán ni castellano, aunque evidentemente se describe o se habla sobre él en nuestro metalenguaje que es el castellano: el hecho no es lingüístico, sino del mundo real, si bien es obvio que precisamos un lenguaje si queremos hablar sobre él.) Lo que dice nuestra afirmación metalingüística es que determinado enunciado (alemán) *corresponde a determinado hecho* (un hecho no lingüístico, un hecho del mundo real) en condiciones enunciadas con precisión.

Naturalmente, podemos reemplazar el lenguaje objeto alemán por cualquier otro —incluso por el castellano—. Por tanto, podemos hacer la afirmación metalingüística:

El enunciado castellano "La hierba es verde" corresponde a los hechos si, y sólo si, la hierba es verde.

Esto parece aún más trivial, pero difícilmente se puede negar, como tampoco se puede negar que expresa en qué condiciones un enunciado corresponde a los hechos.

Hablando en general, sea "*E*" el *nombre* (metalingüístico) de un enunciado del lenguaje objeto y sea "*h*" la *abreviatura* de una expresión del metalenguaje que describa el hecho (supuesto) *H* que describe *E*. Entonces podemos hacer la siguiente afirmación metalingüística:

Un enunciado *E* del lenguaje objeto corresponde a los hechos si, y sólo si, *h*. (*O*... si de hecho *h*.)

Nótese que mientras que aquí "*E*" es el nombre metalingüístico de un enunciado, "*h*" no es un nombre, sino una abreviatura de una expresión del metalenguaje que describe un hecho determinado (el hecho que podemos denominar "*H*").

Podemos decir ahora que lo que hizo Tarski fue descubrir que para hablar sobre la correspondencia entre un enunciado *E* y un hecho *H* precisamos un lenguaje (un metalenguaje) en el que podamos *hablar sobre* el enunciado *E* y *enunciar* el hecho *H*. (Hablamos sobre el primero utilizando el *nombre* "*E*" y sobre el segundo, mediante la expresión metalingüística "*h*" que *enuncia o describe H*.)

La importancia de este descubrimiento estriba en disipar toda duda acerca de la significatividad de hablar sobre la correspondencia de un enunciado con un hecho o hechos.

Una vez hecho esto, podemos sustituir, naturalmente, las palabras "corresponde a los hechos" por las palabras "es verdadero".

A parte de esto, Tarski introdujo un método para *definir* la verdad (en el sentido de la teoría de la correspondencia) en cualquier *sistema formaliza-*

do consistente. Pero no creo que sea éste su logro fundamental. Su mayor éxito es haber rehabilitado la posibilidad de hablar acerca de la correspondencia (y la verdad). Dicho sea de paso, mostró bajo qué circunstancias puede desembocar en paradoja esa forma de hablar y de qué modo podemos evitarlas; *también mostró de qué manera podemos evitar las paradojas y cómo las evitamos de hecho, cuando hablamos en el lenguaje ordinario acerca de la verdad.*

Una vez demostrada la posibilidad de utilizar la "verdad" en el sentido de la correspondencia de enunciados con hechos, no hay nada importante que añadir sobre la palabra "verdad". No hay duda de que lo que ordinariamente llamamos "verdad" es la correspondencia con los hechos; en el lenguaje ordinario llamamos "verdad" a la correspondencia más bien que a la coherencia o a la utilidad práctica. El juez que exhorta al testigo a que diga la verdad y nada más que la verdad, no le exhorta a decir lo que considera útil, sea para sí mismo, sea para alguien más. El juez exhorta al testigo a decir toda la verdad y nada más que la verdad, pero no dice: "Lo que queremos de usted es que no incurra en contradicciones", que es lo de debiera decir si creyese en la teoría de la coherencia.

En otras palabras, no cabe duda de que es la correspondencia el sentido ordinario de "verdad" tal como se emplea en el juzgado. Pero lo que realmente me interesa decir aquí es que eso se puede considerar como una reflexión retrospectiva sin importancia, pues si alguien tuviese ganas de decir, "No, en el lenguaje ordinario, 'verdad' se usa en un sentido diferente", no disputaría con él. Sugiero que nos olvidemos de la terminología: estoy dispuesto a emplear la terminología de mi oponente señalando, no obstante, que tenemos a nuestra disposición *por lo menos* estos tres significados: esto es lo único sobre lo que estaría dispuesto a discutir, aunque me niegue a discutir sobre palabras.

Tendría que señalar, sin embargo, que la teoría de la verdad como correspondencia es una teoría realista, es decir, establece una distinción, que es realista, entre una teoría y los hechos que describe esa teoría, lo que nos permite decir que una teoría es verdadera, falsa o que corresponde a los hechos, relacionando así la teoría con los hechos. Nos permite hablar de una realidad distinta de la teoría, lo cual es algo fundamental, el punto básico, para un realista. El realista quiere disponer de una teoría y de la realidad o los hechos (no lo llamen "realidad" si no les gusta, llámenlo simplemente "los hechos") que son distintos de su teoría *acerca de* los hechos, teoría que puede, de un modo u otro, comparar con los hechos para ver si corresponde a ellos o no. Naturalmente, la comparación es siempre extremadamente difícil.

Una última palabra sobre la teoría de Tarski. A menudo, la finalidad de la misma se interpreta mal: se supone equivocadamente que pretende suministrar un *criterio de verdad,* puesto que esa era la pretensión de las teorías de la coherencia y de la utilidad pragmática. Dichas teorías corroboran la opinión tradicional de que una teoría seria sobre la verdad debería

suministrarnos un *método para decidir* si un enunciado determinado es o no verdadero.

Con su teoría de la verdad, Tarski ha demostrado muchas cosas. Entre otras, ha demostrado que en un lenguaje suficientemente potente (en todo lenguaje en que sea posible formular teorías matemáticas o físicas) no puede haber un criterio de verdad; es decir, un criterio de correspondencia: no se puede decidir, en general, el problema de la verdad de una proposición en lenguajes en los que se pueda formar el concepto de verdad. Por tanto, el concepto de verdad desempeña básicamente el papel de una idea reguladora. En nuestra búsqueda de la verdad, nos asiste el saber que hay algo así como verdad o correspondencia. No nos suministra medios para dar con ella ni para tener la seguridad de haber dado con ella, aunque de hecho sea así. Por tanto, no hay criterio de verdad y no hemos de preguntar por él. Hemos de contentarnos con el hecho de que la idea de verdad como correspondencia con los hechos ha sido rehabilitada. Eso es lo que hizo Tarski y con ello ha prestado un servicio incalculable a la perspectiva realista.

Aunque no poseamos un criterio de verdad, ni siquiera medios para estar totalmente seguros de la falsedad de una teoría, es más fácil descubrir que una teoría es falsa que descubrir que es verdadera (como he explicado detalladamente en otra parte). Incluso tenemos buenas razones para pensar que la mayoría de nuestras teorías —incluso las mejores— son, estrictamente hablando, falsas, pues idealizan o simplifican excesivamente los hechos. Sin embargo, una conjetura falsa puede estar más o menos próxima a la verdad. Así, llegamos a la idea de proximidad a la verdad o de mejor o peor aproximación a la verdad, es decir, a la idea de "*verosimilitud*". He mostrado que esta idea se puede rehabilitar de un modo semejante a como Tarski rehabilitó la idea de verdad como correspondencia con los hechos [13].

Para ello, he recurrido básicamente a las dos ideas de Tarski allí mencionadas: la idea de verdad y la de consecuencia lógica o, más exactamente, el conjunto de las consecuencias lógicas de una conjetura o el contenido de una conjetura.

Al incorporar a la lógica la idea de verosimilitud o aproximación a la verdad, hacemos la lógica aún más "realista", pues se puede emplear ya para hablar sobre el modo en que una teoría corresponde mejor que otra a los hechos —los hechos del mundo real.

Resumiendo: como realista considero la lógica como el *organon de la crítica* (más bien que de la demostración) en nuestra búsqueda de teorías verdaderas y altamente informativas o, al menos, de nuevas teorías que contengan más información y correspondan mejor a los hechos que nuestras viejas teorías. Además, considero la crítica, a su vez, como nuestro principal instrumento para promover el desarrollo del conocimiento del mundo de los hechos.

---

[13] K. R. Popper, *Conjectures and Refutations*, 1963, 1972; Capítulo 10 y Apéndice [véase la traducción citada en la nota 3].

9. COMENTARIOS FILOSOFICOS EN TORNO A LA TEORIA DE LA VERDAD EN TARSKI*

I

Nuestra tarea fundamental, tanto en la ciencia como en la filosofía es, o debiera ser, la búsqueda de la verdad, valiéndonos de conjeturas audaces y de una indagación crítica acerca de los elementos falsos que contienen nuestras diversas teorías rivales [1].

Esta era mi opinión hace treinta y siete años, en julio de 1934, cuando me encontré por primera vez con Alfred Tarski en unas reuniones organizadas en Praga por el círculo de Viena. No obstante, he de subrayar que en aquella época, antes de que Tarski me hubiese enseñado su teoría de la verdad, no era en absoluto clara la suposición, según la cual, nuestra tarea fundamental es la búsqueda de la verdad. En mi libro *Logik der Forschung* (1934), cuyas pruebas de imprenta tenía conmigo en Praga, circunstancia que aproveché para enseñárselas a Tarski (aunque dudo que le interesasen), había escrito: "el esforzarse por el conocimiento y la búsqueda de la verdad siguen constituyendo los motivos más fuertes de la investigación científica" [2]. Con todo, yo no estaba satisfecho con la idea de verdad, por lo que dedicaba toda una sección del libro a intentar defender la idea de verdad como algo inocuo y de sentido común, diciendo que, si quisiéramos, podríamos evitar su uso en metodología de la ciencia, hablando en su lugar de deductibilidad y otras relaciones lógicas similares [3].

---

* Basado en una charla dada en un Simposium en honor de Alfred Tarski con ocasión de su 70 aniversario, que tuvo lugar en la Universidad de California del 23 al 30 de junio de 1971.

[1] En la última sección de este artículo se desarrolla algo más esta formulación relativa a la tarea fundamental de las ciencias para el caso de las ciencias naturales. Aquí habrá que decir algo sobre la terminología.

No le veo ningún interés al problema (fundamentalmente verbal) de si hemos de hablar de "oraciones", "enunciados" o "proposiciones"; los principales críticos de la terminología tarskiana dicen que sus "oraciones" son secuencias de palabras sin interpretar que siguen determinadas reglas gramaticales y que, por tanto, no pueden ser verdaderas o falsas. Pasan por alto el hecho de que Tarski habla explícitamente de "*oraciones significativas*" y sólo de *lenguajes interpretados*. Para demostrar mi desprecio por este tipo de crítica verbal, me he limitado a adoptar la terminología de mis oponentes, de manera que a lo largo del artículo hablaré de "enunciados" más bien que de "oraciones". Por tanto, empleo "enunciado" como sinónimo de una oración significativa interpretada o proposición.

[2] K. R. Popper, *The Logic of Scientific Discovery*, sección 85, pág. 278. [Traducción castellana de Víctor Sánchez de Zavala, *La Lógica de la Investigación Científica*, Madrid, Tecnos, 1962, pág. 259.]

[3] *Op. cit.*, sección 84.

Como es natural, la idea de verdad me resultaba incómoda, pues, durante algún tiempo, ha habido filósofos que la han atacado con buenos argumentos. Lo que me asustaba no era tanto la paradoja del mentiroso, cuanto la dificultad de explicar la teoría de la correspondencia: ¿en qué podría consistir la correspondencia de un enunciado con los hechos? Además, había un punto de vista que, aunque yo nunca había sostenido decididamente, me sentía incapaz de atacar de manera efectiva. Me refiero al punto de vista según el cual, para poder hablar de verdad, hemos de ser capaces de dar un criterio de verdad. *Sostenía* que, a pesar de todo, era legítimo hablar de la verdad, si bien no era capaz de defender mi punto de vista, según el cual la ausencia de un criterio de verdad no podía emplearse como argumento en contra de la legitimidad lógica de la idea de verdad.

Me alegro de no haber expresado nunca por escrito esta peculiar insatisfacción que, como todo el mundo sabe hoy día [4], estaba completamente injustificada. Como ahora sabemos, la ideea de verdad no es en absoluto la única cuya importancia y legitimidad no se ven disminuidas por el hecho de que no haya criterios generales de aplicación en casos específicos. Un ejemplo de características similares nos lo suministra la idea de deductibilidad: sabemos que en muchas teorías es insoluble el problema de la decisión acerca del carácter de teoremicidad. A menos que nos limitemos a una teoría decidible, una teoría en la que se pueda resolver el problema de la decisión, no hay criterios para un procedimiento general que nos permita decidir, en cada caso particular, si un supuesto teorema de la teoría es o no un *teorema válido;* es decir, si es o no deductible con los medios lógicos suministrados por la teoría (en este sentido empleo los términos "teorema válido", "derivación válida", etc.).

Así pues, aunque no dispongamos de un criterio general de validez o teoremicidad para teorías indecidibles, las ideas de validez o teoremicidad son perfectamente claras, incluso en teorías indecidibles: un supuesto teorema es efectivamente válido si, y sólo si, existe una derivación válida del mismo, independientemente de que demos con ella. La ausencia de dicho criterio no contribuye en absoluto a la vaguedad de la expresión "teorema válido". Lo que ocurre es más bien una consecuencia directa de nuestra incapacidad de comprobar el número infinito de todas las derivaciones válidas para ver si alguna de ellas desemboca en el supuesto teorema. Podemos tener suerte y descubrir una prueba o una contraprueba del teorema supuesto; pero si no tenemos esa suerte, a menos que la teoría permita un procedimiento de decisión, no hay manera de descubrir si la fórmula en cuestión es o no un teorema.

Hoy día, esto es casi demasiado trivial como para que valga la pena insistir en ello. Sin embargo, sigue habiendo muchos filósofos que creen que un concepto, por ejemplo, el de verdad, sólo es lógicamente legítimo si hay un criterio que nos permita decidir cuándo un objeto cae bajo dicho con-

---

[4] Véase especialmente la nota 1 de la página 254 del libro de Tarski, *Logic, Semantics, Metamathematics,* Clarendon Press, Oxford, 1956.

cepto. Así, en el volumen 3 de la *Encyclopedia of Philosophy* de 1967 hay un artículo [5] en el que mi opinión de que no hay criterio general de verdad para las teorías científicas se resume en una frase, breve aunque errónea, que me atribuye la opinión según la cual "La propia verdad no es sino una ilusión". En el volumen 2 de dicha *Encyclopedia* se nos dice que en los últimos escritos de Wittgenstein está implícito "que un concepto es vacío si no hay criterios para aplicarlo" [6].

El término "positivismo" tiene muchos significados, pero creo que esta tesis (wittgensteiniana), según la cual "un concepto es vacío si no hay criterios para aplicarlo", expresa el meollo mismo de las tendencias positivistas. (Idea que está muy próxima a Hume.) Si se acepta esta interpretación del positivismo, entonces el positivismo se ve refutado por el desarrollo moderno de la lógica, especialmente por la teoría de la verdad de Tarski, que incluye el *teorema*: para lenguajes suficientemente ricos, no puede haber un criterio general de verdad.

Naturalmente, este teorema es del mayor interés, si recordamos el conflicto clásico entre los estoicos (y luego los cartesianos) y los escépticos. Nos encontramos aquí con uno de esos casos poco frecuentes en los que podemos decir que un conflicto filosófico clásico ha sido resuelto mediante un teorema lógico o metalógico. Mas no se puede decir que los filósofos conozcan o aprecien de manera general este ejemplo.

Con todo, no trato de polemizar aquí con los filósofos que niegan que la teoría de la verdad de Tarski tenga un alcance filosófico. Por el contrario, lo que deseo es evocar mi intensa alegría y alivio cuando en 1935 supe que de la teoría de la verdad de Tarski se seguían las consecuencias siguientes:

(1) Que dicho concepto era definible en términos lógicos que nadie había puesto antes en cuestión *y que, por tanto, era lógicamente legítimo.*

(2) Que era aplicable a cualquier enunciado (cerrado) formulado sin ambigüedad (en un lenguaje no universalístico), suponiendo que no fuese aplicable a su negación *y, por tanto, no vacío, como es obvio,* a pesar del hecho de

(3) que no estuviese ligado a un criterio general, aún cuando todo enunciado derivable de un enunciado o teoría verdadera fuese demostrablemente verdadero;

(4) que la clase de los enunciados verdaderos formaba un sistema deductivo y

(5) que dicho sistema deductivo era indecidible siempre y cuando el lenguaje en cuestión fuese suficientemente rico. (En relación con este resultado, Tarski hacía alusión a Gödel.)

---

[5] *The Encyclopedia of Philosophy*, ed. Paul Edwards, Macmillan, 1967, vol. 3, página 37.

[6] *Op. cit.*, vol. 2, pág. 260. Véase mi libro *Open Society*, II, 4.ª ed., Addendum 1, sección 3. [Traducción castellana de Eduardo Loedel, *La Sociedad Abierta y sus Enemigos*, Buenos Aires, Paidos, 1957.]

Como ya he dicho, vi a Tarski por primera vez en Praga, en julio de 1934. Pero al año siguiente me encontré de nuevo con él en Viena en el Coloquio Karl Menger del que eran miembros Tarski y Gödel y donde también me encontré con personajes de la importancia de Skolem y Abraham Wald. Fue por aquellos días cuando le pedí a Tarski que me explicase su teoría de la verdad, cosa que hizo mediante una conferencia de cerca de veinte minutos en un banco (un banco inolvidable) del *Volksgarten* de Viena. También me permitió ver la serie de pruebas de imprenta de la traducción alemana de su gran artículo sobre el concepto de verdad que acababa de enviarle el editor de *Studia Philosophica*. No hay palabras capaces de describir lo mucho que aprendí con todo esto ni tengo palabras para expresar mi gratitud. Aunque Tarski era tan solo un poco mayor que yo y aunque en aquella época manteníamos relaciones de estrecha amistad, le consideraba como la única persona a quien podía considerar como mi maestro en filosofía. Nunca nadie me ha enseñado tantas cosas.

Sin embargo, hay puntos periféricos en los que quizá esté en desacuerdo con él. Siempre fui un filósofo y un realista de sentido común [7]. Según mi actitud, era de sentido común mantener que a veces el sentido común está equivocado —quizás con más frecuencia que lo contrario—, si bien es evidente que en filosofía tenemos que partir del sentido común, aunque sólo sea para descubrir mediante la crítica en qué está equivocado. Me interesaba por el mundo real, por el cosmos y, por ello, me oponía a todo idealismo, positivismo o incluso neutralismo filosóficos. No me interesaría por la filosofía si no hubiese un mundo real tan rico, incluso mucho más rico que el mundo que conocemos tan superficialmente por nuestra vida diaria, y si la tarea fundamental de la filosofía no fuese el estudio de dicho mundo. Nunca pude saber con exactitud cuál era la actitud de Tarski hacia el realismo. Parecía impresionado por el "reismo" de Kotarbinski, así como por el positivismo de Viena y hacía hincapié en la neutralidad de su concepto de verdad.

Puesto que yo era un realista de sentido común consciente del hecho de que, por tanto, sostenía una teoría "metafísica" [8], estaba muy interesado

---

[7] Soy realista en dos sentidos de la palabra. En primer lugar, creo en la realidad del mundo físico. En segundo lugar, creo que el mundo de las entidades teóricas es real, como ya he explicado en mis artículos "Epistemología sin Sujeto Cognoscente", "Sobre la Teoría de la Mente Objetiva" y "Un enfoque Realista de la Lógica, la Física y la Historia" (ahora, capítulos 3, 4 y 8 de este volumen). En ellos sostengo mi oposición al esencialismo —la realidad de los *conceptos*—, aunque afirmo la realidad de los *problemas, teorías, errores,* etc.

(Por lo que respecta al primer sentido, puedo considerarme materialista en la medida en que creo en la realidad de la materia, aunque *no* soy en absoluto materialista en el sentido en que el "materialismo" representa la opinión (tan extendida) de que la materia es algo último o irreductible, o que es lo único real. Por el contrario, creo que puede haber una teoría verdadera de la materia que explique la extensión de la materia mediante la energía; por ejemplo, mediante las fuerzas, como sugirieron por primera vez Leibniz, Boscovic y Kant.)

[8] Cf. mi libro *Logic of Scientific Discovery*, pág. 252, texto a que corresponde la nota * 1 [trad. cit., pág. 235].

por un aspecto de la teoría de la verdad de Tarski que me parecía realista, aspecto cuya mera existencia sospecho que él habría negado [9].

La teoría de Tarski, como todos ustedes saben y como él mismo fue el primero en subrayar, es una *rehabilitación* y reelaboración de la teoría clásica de que la verdad es la correspondencia con los hechos y a *mí* me parece que esto apoya el realismo. A la vez, la teoría de Tarski rehabilita y reelabora también algunas de las *críticas* clásicas a esta teoría de la correspondencia, pues indica hasta qué punto están en lo cierto quienes sospechan que la teoría de la correspondencia es paradójica. Esta última parte la resuelve esencialmente la teoría de Tarski, según la cual la semática (L_1) de un lenguaje objeto (L_0) —esto es, el metalenguaje que contiene la idea de "verdadero en (L_0)" como concepto definible— ha de ser *esencialmente más rica* (de un orden superior) que el lenguaje objeto (L_0).

Como sabemos, el lenguaje objeto L_0 puede contener su propia sintaxis y, en especial, nombres que describan sus propias expresiones. Mas L_0 no puede contener, sin correr el riesgo de incurrir en antinomias, términos semánticos como *denotación, satisfacción o verdad,* es decir, nociones que relacionen los *nombres de las expresiones* de L_0 con los *hechos u objetos* a que se refieren dichas expresiones.

Todo esto me suministró una materia de reflexión que desarrollé a lo largo de muchos años. Contaré brevemente algunas de esas reflexiones.

II

Si, como sugiere la teoría de Tarski, la verdad es la correspondencia con los hechos, entonces abandonemos totalmente, por el momento, la palabra "verdad" y hablemos en su lugar de "la correspondencia de los enunciados con los hechos que describen".

Creo que la aparente imposibilidad de descubrir o explicar esta correspondencia fue la causante de que todas las teorías de la correspondencia anteriores a Tarski pareciesen sospechosas, incluso a quienes como yo apreciaban la teoría de la correspondencia simplemente por su carácter realista y de sentido común [10].

Seamos ahora audaces y tomemos en serio el que haya enunciados que correspondan a los hechos. La teoría que trate esta situación ha de poder hablar (1) de los enunciados de un lenguaje que denominamos lenguaje objeto o lenguaje que se investiga [11] y (2) de hechos o hechos supuestos.

---

[9] Cf. A. Tarski, "The Semantic Conception of Truth and the Foundations of Semantics", *Philosophy and Phenomenological Research, 4,* 1944, págs. 341-76; véase especialmente la sección 19. [Véase la trad. cit. en la nota 1, pág. 302.]

[10] Para detalles, véase *C. & R.,* pág. 223. [Traducción castellana de Néstor Míguez, *El Desarrollo del Conocimiento Científico. Conjeturas y Refutaciones,* Buenos Aires, Paidos, 1967, pág. 259.]

[11] Parece ser que la expresión "lenguaje objeto" se introdujo originalmente en el sentido de "lenguaje que habla acerca de objetos (físicos)". Yo empleo esta expresión

(1) Para hablar de enunciados, hemos de disponer de *nombres* de enunciados, por ejemplo, *nombres citados o nombres que describen* enunciados. Esto significa que una teoría de la correspondencia ha de formularse en un metalenguaje; es decir, un lenguaje en que se pueda discutir o hablar acerca de las expresiones del lenguaje objeto que se investiga.

(2) Para hablar de la relación entre el enunciado y los hechos, hemos de disponer de descripciones de hechos; es decir, hemos de poder describir en nuestro metalenguaje todos aquellos hechos que podemos describir en el lenguaje objeto. Por tanto, el metalenguaje ha de poseer traducciones de los enunciados del lenguaje objeto o ha de contener el lenguaje objeto como una *parte de sí mismo* (método que evita el desagradable problema de la existencia de traducciones fieles).

Así, nos encontramos con que una teoría que se ocupe de la correspondencia entre enunciados y hechos y, por tanto, de una relación entre ellos, ha de ser formulada en un metalenguaje que, aparte de las palabras lógicas usuales, tenga a su disposición tres tipos de expresiones:

(1) Nombres de enunciados; es decir, de expresiones lingüísticas de un lenguaje objeto que forman parte de la "morfología" o "sintaxis" de ese lenguaje objeto.

(2) Enunciados que describen los hechos (incluso los no-hechos) discutidos en ese lenguaje objeto; es decir, traduciones del lenguaje objeto al metalenguaje. (Para evitar los errores de traducción, el lenguaje objeto ha de pasar a formar parte del metalenguaje, como ya insinuábamos.)

(3) Además de estos dos tipos fundamentales de expresiones, hay un tercer tipo: los términos que denotan predicados o relaciones entre ambos tipos fundamentales de expresiones; por ejemplo, predicados tales como "*X* corresponde a los hechos" o relaciones tales como "*X* corresponde a los hechos si, y sólo si, *y*". (Este último tipo de términos es semántico y de orden superior al lenguaje objeto a que se refiere.)

Estos tres requisitos constituyen requisitos mínimos casi obvios de todo lenguaje en que podamos formular una teoría de la correspondencia.

Tarski denominaba "metalenguaje semántico" al lenguaje que reuniese estos tres requisitos mínimos.

La grandeza y audacia del logro de Tarski se ve por el hecho de que descubrió estos requisitos mínimos y porque además descubrió que los predicados o relaciones mencionados en el apartado (3), que relacionan las expresiones con el mundo de los hechos, estaban necesariamente fuera del alcance de los medios de que disponemos en el lenguaje objeto [12].

Es evidente que una vez que disponemos de las tres categorías de ex-

---

en el sentido de "lenguaje que es objeto de investigación"; es investigado mediante una teoría formulada en un metalenguaje. (Naturalmente, esto plantea la idea de una jerarquía infinita de metalenguajes.)

[12] Un resultado sólo ligeramente menos importante desde el punto de vista filosófico, relativo a los términos mencionados en el apartado (3) es que, *qua* términos

presiones, podemos formular en el metalenguaje afirmaciones tales como

P corresponde a los hechos si, y sólo si, *P*,

suponiendo que las letras mayúsculas como "*P*" sean variables que están en lugar de *nombres* metalingüísticos de esos enunciados del lenguaje objeto que describen hechos cuyas *traducciones* metalingüísticas se representan mediante las minúsculas correspondientes, como "*p*".

Al estudiar la teoría de la verdad de Tarski, me di cuenta de que las cosas me resultaban más fáciles, a mí y a algunos de mis estudiantes, si hablaba de este modo acerca de la *correspondencia con los hechos* en lugar de hacerlo acerca de la *verdad*. Incidentalmente, también descubrí que las cosas se hacían más fáciles si utilizaba como ejemplos enunciados *falsos* del lenguaje objeto.

Tomemos el alemán como lenguaje objeto y el castellano como metalenguaje y recordemos que la traducción castellana de la frase alemana "*Der Mond besteht aus grünem Käse*" es "La luna está hecha de queso verde". Con estos enunciados falsos, es evidente que podemos constituir la afirmación semántica verdadera:

"El enunciado alemán '*Der Mond besteht aus grünem Käse*' corresponde a los hechos si, y sólo si, la luna está hecha de queso verde".

No obstante, la utilización de un enunciado falso del lenguaje objeto es algo secundario. Por otro lado, para algunos estudiantes resulta realmente útil hablar de correspondencia con los hechos (en lugar de hablar de verdad). Les permite ver con mayor claridad por qué el enunciado que ocupa el lugar de la variable minúscula "*p*" es, y *ha* de ser, un enunciado metalingüístico acerca de un *hecho* (o un hecho supuesto); es decir, la descripción metalingüística de una situación descrita también con el lenguaje objeto.

III

En el segundo párrafo de su segundo artículo sobre la verdad [1], Tarski hace notar que no es preciso utilizar conceptos semánticos para definir la verdad (es decir, conceptos que relacionen las expresiones lingüísticas con las cosas expresadas). Mas, puesto que define "verdad" con ayuda del concepto de satisfacción y puesto que este concepto es claramente semántico (así lo clasifica el propio Tarski en el primer párrafo de su artículo XV, página 401 de *Logic, Semantics, Metamathematics*), hasta el lector más atento ha de ser disculpado si al principio se encuentra algo perplejo. La solución del problema puede ser como sigue. Todo lenguaje suficientemente rico que

---

del lenguaje, poseen el mismo carácter morfológico que los términos mencionados en (1): es decir, pertenecen a la morfología desarrollada *en* el metalenguaje (aunque no a aquella parte del mismo que contiene la morfología o sintaxis *del* lenguaje objeto y que se puede desarrollar en el propio lenguaje objeto).

[1] Cf. la página 152 de la traducción inglesa de Woodger, *Logic, Semantics, Metamathematics*, Clarendon Press, Oxford, 1956.

hable acerca de algún tema puede contener (según los resultados alcanzados independientemente por Tarski y Gödel) su propia "morfología" o "sintaxis", mientras que (como ha mostrado Tarski) ningún lenguaje consistente puede contener los medios necesarios para definir su propia semántica. Como hemos visto, lo que precisa Tarski para dar su definición es un metalenguaje semántico de orden superior al del lenguaje objeto, cuya semántica contiene. Pero los términos que son semánticos respecto del lenguaje *objeto* pueden ocupar, dentro del metalenguaje como tal, la misma posición que sus otros términos morfológicos o sintácticos. Por tanto, la semántica de un lenguaje objeto L_n puede formar parte de la sintaxis del metalenguaje de orden superior (digamos, L_{n+1}): en L_{n+1} no es necesario que haya ningún término de carácter no-morfológico o no-sintáctico. Esto equivale a reducir la semántica de L_n a la sintaxis de L_{n+1}

Este punto tiene un interés filosófico general, no sólo porque los términos semánticos eran tenidos por sospechosos, sino también porque reducir términos de índole sospechosa a términos de índole aceptable es algo que merece nuestra atención. En todo caso, el éxito de Tarski al reducir los términos pertenecientes a la semántica de L_n a términos no-semánticos de L_{n+1} elimina todo asomo de duda.

Admito la importancia de esta reducción, pues es algo poco frecuente en filosofía poder introducir una categoría de términos totalmente nueva (y sospechosa) basándose en categorías establecidas (no sospechosas); se trata de una rehabilitación, de salvar el honor de un término sospechoso.

Por otro lado, considero que no poseen una importancia filosófica particular las definiciones y las cuestiones de reductibilidad. Si no podemos definir un término, nada nos impide utilizarlo como término indefinido: la utilización de términos indefinidos no sólo es legítima, sino también inevitable, pues todo término definido, en última instancia, ha de ser definido con ayuda de términos indefinidos [14]. En mi opinión, lo que confiere especial importancia filosófica a la obra de Tarski *no es su descripción efectiva de un método para definir "verdad"*, sino *su rehabilitación de la teoría de la verdad como correspondencia* y la demostración de que ya no nos acechan más dificultades, una vez que hemos comprendido la necesidad esencial de recurrir a un metalenguaje más rico que el lenguaje objeto y su sintaxis. Está bastante claro que, si queremos, podemos partir de términos semánticos primitivos (como hizo R. M. Martin) [15] en vez de evitarlos cuidadosamente. Llegaríamos esencialmente a la misma teoría semántica de la verdad o correspondencia con los hechos. Mas sin la teoría de Tarski, que suministra un metalenguaje semántico libre de términos específicamente semánticos, no se hubieran superado las sospechas de los filósofos hacia los términos semánticos.

[14] Así, Tarski ha subrayado que el concepto de verdad puede introducirse mediante axiomas más bien que mediante definiciones.

[15] Cf. R. M. Martin, *Truth and Denotation, A Study in Semantical Theory*, Routledge and Kegan Paul, Londres, 1958. [Traducción castellana de C. Piera Gil y V. Sánchez de Zavala, *Verdad y Denotación*, Madrid, Tecnos, 1962.]

IV

Como ya he dicho, soy realista. Admito que se pueda defender un idealismo como el de Kant en la medida en que afirma que *todas nuestras teorías son un producto humano* que tratamos de imponer al mundo natural. Pero soy realista porque sostengo que el problema de si son verdaderas o no las teorías hechas por el hombre depende de los hechos reales, los cuales no son en absoluto un producto humano, salvo algunas excepciones. Las *teorías* hechas por el hombre pueden chocar con esos *hechos* reales, por lo que en nuestra búsqueda de la verdad podemos vernos obligados a reajustar nuestras teorías o a abandonarlas.

La teoría de Tarski nos permite *definir la verdad* como correspondencia con los hechos, pero también podemos utilizarla para definir la *realidad* como aquello a lo que corresponde un enunciado verdadero. Por ejemplo, podemos distinguir entre *hechos reales,* es decir, hechos (supuestos) que son reales, y *hechos (supuestos) que no son reales* (esto es, no-hechos). O, para decirlo de un modo más explícito, podemos decir que un hecho supuesto, como el que la luna esté hecha de queso verde, es real si, y sólo si, es *verdadero* el enunciado que lo describe —en este caso, "La luna está hecha de queso verde"—; de lo contrario, el hecho supuesto no es real (o, si ustedes prefieren, no es un hecho en absoluto).

Del mismo modo que Tarski nos permite sustituir el término "verdad" por "el conjunto de enunciados (u oraciones) verdaderos", podemos sustituir el término "realidad" por "el conjunto de hechos reales".

Por tanto, considero que si podemos definir el concepto de verdad, también podemos definir el concepto de realidad. (Evidentemente, surgen aquí problemas de orden análogos a los problemas de orden de los lenguajes en la obra de Tarski: Véase especialmente su *Post-Scriptum,* págs. 268-77, al libro *Logic, Semantics, Metamathematics.*) No pretendo que el término "verdad" sea, en uno u otro sentido, más básico que el término "realidad": deseo rechazar tal sugerencia por su sabor idealista [16]. Lo único que digo es que, si es posible definir "verdad" como "correspondencia con los hechos" o, lo que es lo mismo, como "correspondencia con la realidad", entonces también es posible definir "realidad" como "correspondencia con la verdad". Puesto que soy realista, quiero estar seguro de que el concepto de realidad no es "vacío" o sospechoso por alguna otra razón, como tampoco lo es el concepto de verdad.

V

Entre las más antiguas teorías de Tarski, accesibles a un filósofo no demasiado sofisticado, como yo, está el llamado Cálculo de Sistemas. Yo estaba

---

[16] Véase K. R. Popper, *Conjectures and Refutations,* nota 33 de la pág. 116, donde expreso mi reconocimiento a Alexandre Koyré [trad. cit., págs. 137 y sig.].

en París en 1939 cuando, si no recuerdo mal, Tarski completaba su artículo sobre el Cálculo de Sistemas [17]. Me pareció del mayor interés.

He intentado combinar algunos de los resultados más obvios del artículo de Tarski sobre la verdad con su artículo sobre el Cálculo de Sistemas. Obtenemos inmediatamente estos teoremas tan triviales en los que se supone que los lenguajes de los que se habla no son universalísticos.

Teorema. El conjunto T de enunciados verdaderos de un lenguaje es un sistema deductivo en el sentido del Cálculo de Sistemas de Tarski. Es completo [18].

En cuanto sistema deductivo, T es una clase consecuente; es decir, es idéntico a la clase $Cn(T)$ de sus propias consecuencias lógicas ($T = Cn(T)$). Es un sistema completo en el sentido de que si añadimos a T un enunciado que no pertenece a T, la clase resultante es inconsistente.

Teorema. El conjunto de enunciados verdaderos de un lenguaje suficientemente rico es un sistema deductivo no-axiomatizable en el sentido del Cálculo de Sistemas de Tarski.

Ambos teoremas son perfectamente triviales y en lo que sigue supondremos que los lenguajes en cuestión son lo suficientemente ricos para satisfacer el segundo de estos teoremas.

Introduzco ahora un concepto nuevo, el de *contenido de verdad* de un enunciado *a*.

Definición. El conjunto de todos los enunciados verdaderos que se siguen de un enunciado determinado *a* se denomina *contenido de verdad de a*. Es un sistema deductivo.

Teorema. El contenido de verdad de un enunciado verdadero *a* es un sistema axiomatizable $A_T = A$; el contenido de verdad de un enunciado falso *a* es el sistema deductivo $A_T \subset A$ donde A_T no es axiomatizable, suponiendo que el lenguaje objeto en cuestión sea suficientemente rico.

Esta definición y su teorema se pueden generalizar: el cálculo tarskiano de sistemas deductivos puede considerarse una generalización del cálculo de enunciados, pues a todo enunciado (o clase de enunciados lógicamente equivalentes), *a*, le corresponde un sistema (finitamente) *axiomatizable*, *A*, tal que

$$A = Cn(A) = Cn(\{a\});$$

y *viceversa*: a todo sistema deductivo *axiomatizable*, *A*, le corresponde un enunciado (o una clase de enunciados lógicamente equivalentes), *a*. Mas,

---

[17] Véase A. Tarski, *op. cit.*, págs. 342-83.

[18] Sigo básicamente el simbolismo de Tarski (especialmente en la utilización de mayúsculas para representar sistemas deductivos), excepto en que represento la clase de los enunciados verdaderos con "T", mientras que Tarski utiliza "Tr".

puesto que también hay sistemas deductivos o clases consecuentes que no son axiomatizables, de modo que no hay un enunciado o clases finita de enunciados del que sean clase consecuente, podemos considerar como generalización el tránsito de enunciados a clases consecuentes o sistemas deductivos, o del cálculo de enunciados al Cálculo de Sistemas.

Así, en general, para toda clase consecuente o sistema deductivo A tenemos un sistema A_T, el contenido de verdad de A, que es idéntico a A si, y sólo si, A consta solamente de enunciados verdaderos y que, en todo caso, es un subsistema de A: es la clase producto o intersección de los conjuntos A y T.

Surge el problema de si tenemos también lo que podríamos denominar el *contenido de falsedad* A_F de a o de A, correspondiente al *contenido de verdad* A_T de a o de A. Una consecuencia obvia que se presenta por sí misma es definir la clase de todos los enunciados falsos pertenecientes al sistema deductivo A como el contenido de falsedad de A. Con todo, esta sugerencia no es plenamente satisfactoria, si empleamos (como recomiendo) el término "*contenido*" como un tercer sinónimo de "sistema deductivo" o "clase consecuente"; pues esta clase, que suponemos que sólo consta de enunciados falsos, no es un sistema deductivo: todo sistema deductivo A contiene enunciados verdaderos —de hecho en un número infinito— con lo que la clase que conste solamente de enunciados falsos pertenecientes a A no puede ser un contenido.

Para introducir la idea de contenido de falsedad A_F de un enunciado a o clase consecuente A, hay que retrotraerse a la idea de *contenido relativo* de A, dado B, que podemos introducir como generalización de un sistema deductivo tarskiano o *contenido* (*absoluto*), $A = Cn(A)$. Explicaré esta idea y, previendo algunas posibles críticas intuitivas, también introduciré la idea de *medida de contenido*. Finalmente, introduciré la idea de aproximación a la verdad o verosimilitud con ayuda de la idea de medida del contenido de verdad y falsedad.

VI

Tarski habla de clases consecuentes o sistemas deductivos mayores o menores. Realmente, el conjunto de sistemas deductivos (de un lenguaje dado) está parcialmente ordenado por la relación de inclusión que coincide con la relación de deductibilidad. Podemos emplear la siguiente consideración, hecha por Tarski en su artículo sobre el Cálculo de Sistemas, como punto de referencia para la relativización de una clase consiguiente, contenido o Sistema deductivo: "...entre los sistemas deductivos hay uno que es el más pequeño, i.e., un sistema que es subsistema de todos los demás sistemas deductivos. Se trata del sistema $Cn(0)$, el conjunto de consecuencias del conjunto vacío. Este sistema, que en aras de la brevedad llamaremos "L", se puede interpretar como el conjunto de todas las oraciones* lógicamente

* *Sentences;* véase la nota 1 al comienzo del artículo. [*N. T.*]

válidas (o, en general, como el conjunto de todas aquellas oraciones que consideramos verdaderas desde el principio, al emprender la construcción de la teoría deductiva objeto de nuestra... investigación)" [19].

Esto sugiere que podemos utilizar un sistema distinto del sistema cero L "como el conjunto de todas aquellas oraciones *que consideramos verdaderas desde el principio,* al emprender la construcción..." Como anteriormente, representemos con la variable "A" el sistema deductivo por cuyo contenido estamos interesados y, con la variable "B", "el conjunto de todas aquellas oraciones que consideramos verdaderas desde el principio". Entonces, podemos escribir

$$Cn(A, B)$$

como relativización del $Cn(A)$ de Tarski, que se convierte en un caso especial cuando $B = L = Cn(0)$:

$$Cn(A) = Cn(A, L).$$

Podemos escribir "A, B" como abreviatura de "$Cn(A, B)$" del mismo modo que Tarski escribe "A" en lugar de "$Cn(A)$". Así pues, el pasaje de Tarski que hemos citado sugiere lo siguiente:

Definición: $A, B = Cn(A, B) = Cn(A+B) - Cn(B)$.

Lo cual, como es obvio, lleva a lo siguiente:

Teorema. $A = Cn(A) = A, L = Cn(A, L) = Cn(A+L) - Cn(L)$.

Limitándonos al modo relativo de escribir tenemos, pues, para el contenido de verdad:

$$A_T = A_T, L = Cn((A.T) + L) - Cn(L)$$

y para el contenido de falsedad:

$$A_F = A, A_T = Cn(A+A_T) - Cn(A_T) = Cn(A) - Cn(A_T),$$

lo cual convierte al contenido de falsedad A_F en un contenido relativo cuya extensión coincide (como sugeríamos originalmente) con la clase de todos los enunciados falsos de A.

VII

Puede plantearse la siguiente objección en contra de la definición propuesta del contenido de falsedad, A_F, como contenido relativo A, A_T. Dicha definición se apoya intuitivamente en una cita de Tarski en la que considera a L el sistema deductivo menor o cero. Mas en nuestra definición

$$A = A, L = Cn(A+L) - Cn(L)$$

[19] A. Tarski, *Logic, Semantics, Metamathematics.* Clarendon Press, Oxford, 1956, página 343.

tomamos demasiado al pie de la letra la palabra cero: ahora comprendemos, a la vista de nuestra expresión "—$Cn(L)$", puesto que la restamos, que había que interpretar L como un conjunto de *medida cero* y no como un conjunto literalmente vacío o que ya no esté presente (de manera que sólo queden los enunciados no lógicos de A, cosa que no pretendíamos).

Tomemos o no en serio esta objección, desaparece en cualquier caso si decidimos operar con una *medida de contenido*, $ct(A)$ o $ct(A, B)$, en lugar de con el contenido o clase consecuente, $Cn(A)$ o $Cn(A, B)$, mismo.

En 1934, Tarski había llamado la atención de la reunión de Praga sobre la axiomatización del cálculo de la probabilidad relativa de un sistema deductivo A, dado un sistema deductivo B, debida a Stephan Mazurkiewicz [20], que se basaba en el Cálculo de Sistemas de Tarski. Puede considerarse que dicha axiomatización introduce una función de medida de los sistemas deductivos o contenidos A, B, C,..., aún cuando dicha función particular, la función de probabilidad,

$$p(A, B)$$

aumente al disminuir el contenido relativo. Esto sugiere la introducción de una medida de contenido, mediante una definición del tipo:

$$\textit{Definición}\text{: } ct(A, B) = 1 - p(A, B)$$

que aumenta y disminuye con el aumento o disminución del contenido relativo. (Evidentemente se pueden dar otras definiciones, pero ésta parece la más simple y obvia.) Inmediatamente tenemos

$$\begin{aligned} ct(L) &= 0 \\ ct(A_T) &= 1 - p(A.\ T,\ L) = 1 - p(A.\ T) \\ ct(A_F) &= 1 - p(A,\ A_T) \end{aligned}$$

que corresponden a los resultados anteriores.

Esto nos sugiere la posibilidad de introducir la idea de *semejanza con la verdad o verosimilitud* de un enunciado a, de manera que aumente con su contenido de verdad y disminuya con su contenido de falsedad. Esto se puede hacer de varios modos [21].

---

[20] Tarski se refería a S. Mazurkiewicz, "Über die Grundlagen der Wahrscheinlichkeitsrechnung I", *Monatshefte f. Math. & Phys., 41,* 1934, págs. 343-52. De la nota 2, en la página 344 de este escrito, se desprende que los matemáticos polacos conocían el Cálculo de Sistemas de Tarski ya en 1930. El sistema de Mazurkiewicz tiene cierto aspecto finitista, frente a mi propio sistema (véase *L. Sc. D.*, págs. 326-58 [trad. cit., págs. 295-333]), que puede recibir diversas interpretaciones, por ejemplo, la de un cálculo de probabilidades de sistemas deductivos.

Tal vez deba mencionar que en este volumen utilizo letras minúsculas como símbolos de las funciones de medida, tales como probabilidad, contenido y verosimilitud: por ejemplo, $p(A)$, $ct(A)$, $vs(A)$; mientras que en el Apéndice de *Conjectures and Refutations,* donde me ocupé por primera vez de las dos últimas funciones de medida, escribía *Ct* y *Vs.*

[21] Cf. K. R. Popper, *Conjectures and Refutations,* Apéndice 3, págs. 391-7 [trad. cit., págs. 450-7].

El más obvio, consiste en tomar $ct(A_T) - ct(A_F)$ como medida de verosimilitud de A. No obstante, por razones que no discutiré aquí, me parece ligeramente preferible definir la verosimilitud, $vs(A)$, mediante el producto de esta diferencia por un factor de normalización, de preferencia al siguiente:

$$1/(p(A_T, L) + p(A, A_T)) = 1/(2 - ct(A_T) - ct(A_F)).$$

De este modo obtenemos la siguiente

Definición: $vs(A) = (ct(A_T) - ct(A_F))/(2 - ct(A_T) - ct(A_F))$.

que, como es natural, también puede escribirse con la notación - p:

$$vs(A) = (p(A, A_T) - p(A_T, L))/(p(A, A_T) + p(A_T, L)).$$

Esto conduce a

$$-1 \leq vs(A) \leq +1,$$

y especialmente a

$$vs(L) = 0;$$

es decir, la verosimilitud no mide esa esperie de aproximación a la verdad conseguida a costa de no decir nada (eso lo mide la falta de contenido o probabilidad), sino la aproximación a "la verdad plena" mediante un contenido de verdad cada vez mayor. Hay dos razones por las que creo que la verosimilitud en este sentido constituye un fin más adecuado para las ciencias —especialmente las naturales— que la verdad. En primer lugar, porque no creemos que L represente el objeto de la ciencia, aunque $L = L_T$. En segundo lugar, porque podemos preferir teorías que consideramos falsas a otras, incluso verdaderas como L, siempre y cuando pensemos que su contenido de verdad supera suficientemente su contenido de falsedad.

En estas últimas secciones, me he limitado a bosquejar un programa consistente en combinar la teoría de la verdad de Tarski con su Cálculo de Sistemas para obtener un concepto de *verosimilitud* que nos permita hablar de *teorías que constituyen mejores o peores aproximaciones a la verdad* sin miedo a decir sinsentidos. Naturalmente, no sugiero la existencia de un criterio de aplicabilidad de esta idea, del mismo modo que tampoco lo hay para la idea de verdad. Pero a algunos de nosotros (por ejemplo, el propio Einstein) a veces nos gustaría decir algo así como que poseemos razones para suponer que la teoría gravitatoria de Einstein *no* es *verdadera*, si bien constituye una *mejor aproximación a la verdad* que la de Newton. Creo que el desideratum fundamental de la metodología de las ciencias naturales es poder decir cosas semejantes sin mala conciencia.

*SUPLEMENTO**

NOTA SOBRE LA DEFINICION DE VERDAD DE TARSKI

En su famoso artículo sobre el Concepto de Verdad [1], Tarski describe un método para definir la idea de verdad o, más exactamente, la idea de "*x* es un enunciado verdadero (del lenguaje *L*)". El método se aplica primeramente al lenguaje del cálculo de clases, aunque también se puede aplicar con toda generalidad a muchos lenguajes (formalizados) distintos, incluyendo lenguajes que permiten formalizar algunas teorías empíricas. Lo característico de este método es que la definición de "enunciado verdadero" se basa en una definición de la *relación de satisfacción* o, más exactamente, de la expresión "la sucesión infinita *f* satisface la función de enunciado *X*" [2]. *Esta relación de satisfacción tiene interés por sí misma,* al margen de su carácter crucial para la definición de verdad (y al margen de que de la definición de satisfacción a la de verdad no haya más que un paso que no plantea mayores problemas). Esta nota trata el problema de intentar utilizar en la definición de satisfacción *sucesiones finitas en lugar de sucesiones infinitas*. Creo que se trata de algo deseable desde el punto de vista de la apli-

---

* Publicado por primera vez en *Mind, 64,* N. S., 1955. Al margen de las consideraciones puestas entre corchetes, algunas cursivas nuevas y algunas pequeñas correcciones de estilo, solamente he introducido los cambios siguientes: siguiendo la traducción de Woodger de 1956, sustituyo "cumplir" y "cumplimento" ["*fulfil*" y "*fulfilment*"], etc., por "satisfacer" y "satisfacción" ["*satisfy*", "*satisfation*"]; consiguientemente, en la Definición 22b sustituyo mi anterior "satisface", que aparecía dos veces, por "se acomoda a" ["*complies with*"]. He cambiado las dos últimas palabras del texto de esta Nota y en lugar de "una sucesión infinita" pongo "sucesiones infinitas". He añadido también las páginas y otras referencias a la traducción de Woodger. [Todos los añadidos están entre corchetes.] Por lo demás he dejado el Suplemento tal como se publicó por primera vez.

[1] Cf. A. Tarski, "Der Wahrheitsbegriff in den formalisierten Sprachen" (*Studia Philosophica,* vol. I, 1935, págs. 261 y sigs.). ["The concept of Truth in Formalized Languages", en A. Tarski, *Logic. Semantics, Metamathematics,* 1956, artículo VIII, páginas 152-278. [El lector castellano podrá leer una exposición más breve de A. Tarski, "La concepción semántica de la verdad y los fundamentos de la semántica", en M. Bunge (ed.) *Antología semántica.* Buenos Aires, Nueva Visión, 1960, págs. 111-157.]] Entiendo que Tarski prefiere traducir "*Aussage*" y "*Aussagefunction*" por "oración" y "función de oración" (mientras que yo empleo aquí, "enunciado" y "función de enunciado"), términos que son los que aparecen en la traducción de los artículos lógicos de Tarski hecha por J. H. Woodger, que pronto aparecerán editados en Clarendon Press, Oxford. [El libro se publicó en 1956. Hay además otras diferencias entre mi tradución y la de Woodger.]

[2] Cf. *Op. cit.,* pág 311 [193], 313 [195]. Nótese que la clase de las funciones de enunciado [o funciones de oración] incluye la de los enunciados; i. e., la de las funciones de enunciado *cerradas.*

cación de la teoría a las ciencias empíricas, así como desde un punto de vista didáctico.

El propio Tarski discute brevemente dos métodos [3] que utilizan sucesiones finitas de diversa longitud en lugar de sucesiones infinitas. Mas señala que dichos métodos alternativos tienen algunas desventajas. Señala que el primero de ellos lleva a "complicaciones considerables" [o "un tanto graves"] (*ziemlich bedeutenden Komplikationem*) en la definición de satisfacción (Definición 22), mientras que el segundo tiene la desventaja de presentar "cierta artificiosidad" (*eine gewisse Künstlichkeit*), por cuanto que lleva a una definición de verdad (Definición 23 [pág. 195]) que utiliza la idea de una "sucesión vacía" o "sucesión de longitud cero" [4]. Lo que deseo señalar en esta nota es que una modificación relativamente pequeña en el procedimiento de Tarski nos permite operar con sucesiones finitas sin que nos veamos envueltos en las complicaciones o arbitrariedades (por ejemplo, sucesiones vacías) en que pensaba Tarski. El método nos permite conservar un procedimiento tan natural como el de la condición (8) de la Definición de Tarski [pág. 193] (evitando así el rodeo de tener que introducir relaciones —o atributos— de grado igual al número de variables *libres* de la función de enunciado en cuestión). La variante que introduzco en el método de Tarski es ligera; pero en vista de que Tarski alude a otras variantes con desventajas considerables, aunque no a ésta, puede que valga la pena describir esta mejora quizá insignificante [5].

Para ello, es útil mencionar de modo informal, en primer lugar, la idea de *lugar número n* (o *n*-avo lugar) de una sucesión finita de cosas y, en segundo lugar, la idea de *longitud* de una sucesión finita *f* —i.e. el número

---

[3] El primer método alternativo aparece bosquejado en la nota 40 de las págs. 309 y siguientes [pág. 91, nota 1] de Tarski. (No se enuncia explícitamente la posibilidad de utilizar este método para evitar las sucesiones infinitas, pero es obvio que se puede emplear en este sentido.) El segundo método se describe en la nota 43 de las págs. 313 y siguiente [pág. 195, nota 1]. El método que Tarski sugiere en esta nota, que difiere del que utiliza en el texto, lo emplea Carnap en su *Introduction to Semantics* (1942), páginas 47 y siguientes [más exactamente, págs. 45-8]. Aunque Carnap alude a Tarski, pasa por alto su anticipación de este método particular. (Hay incluso un tercer método, apuntado por Tarski en la nota 87 de la pág. 386 [pág. 245, nota 2]. Dicho expediente es muy sencillo, aunque sin duda es muy artificial en el sentido que da Tarski a la palabra; además, este método sólo se relaciona con la definición de verdad misma y no con la de cumplimiento [satisfacción] que tiene interés por sí misma.)

[4] También Carnap utiliza este concepto artificial.

[5] La diferencia fundamental entre mi método y los que sugiere Tarski (mencionados en la nota 3 *supra*) consiste en lo siguiente: Tarski sugiere correlacionar con una función dada (sucesiones infinitas) o sucesiones finitas de longitud definida (dependientes de la función), mientras que yo empleo sucesiones finitas "de longitud suficiente" (Definición 22a); i.e., no excesivamente cortas para la función en cuestión. Según esto, mis sucesiones finitas pueden ser de cualquier longitud (superior a cierto límite mínimo que depende de la función). Mas la admisión de sucesiones finitas de cualquier longitud (suponiendo que sea suficiente) no implica ninguna vaguedad, pues es fácil derivar un *teorema* (cf. el Lema A de Tarski, pág. 317 [198]) según el cual, si *f* cumple *x*, entonces toda extensión *g* de *f* también cumple *x* (siendo *g* una extensión de *f* si, y sólo si, para todo f_i hay un g_i tal que $g_i = f_i$). Por tanto, el teorema nos dice que sólo es preciso tomar en consideración aquellas sucesiones finitas *más breves* de cuantas se adecúan a la función considerada (evidentemente, a la función compuesta total en cuestión y no a sus componentes).

de lugares de f (en símbolos, $Np\ (f)$) que es igual a su mayor número de lugar— y la idea de comparación de distintas sucesiones finitas respecto a su longitud. En tercer lugar, mencionaré que una cosa puede ocupar un lugar determinado —digamos, el n-avo— en la sucesión, con lo que se la puede describir como [el individuo n-avo] la cosa n-ava o el miembro n-avo de la sucesión en cuestión. Hay que considerar que una y la misma cosa puede aparecer en diversos lugares de la misma sucesión, así como en distintas sucesiones [6].

Al igual que Tarski, empleo "f_1", "f_2", ... "f_t", f_k", ... "f_n", como nombres de las cosas que ocupan el primer, segundo, i-avo, k-avo... n-avo lugar de la sucesión f. Empleo la misma notación que Tarski, excepto por lo que respecta al uso [por razones tipográficas] de "$P_k y$" como nombre de la universalización [o cuantificación universal] de la expresión y respecto a la variable v_k [7]. Se supone que a la definición (11) [8] de Tarski se añade una definición de "en la función de enunciado x aparece v_k" —suposición que no va en absoluto más allá de los métodos de Tarski y que de hecho está implícita en el propio tratamiento que él hace de la cuestión.

Podemos proceder ahora a sustituir la Definición 22 de Tarski [pág. 193]. La sustituiremos por dos definiciones, la Definición 22a preliminar y la Definición 22b que corresponde a la definición tarskiana propiamente dicha.

Definición 22a

Una sucesión finita de cosas, f, es *adecuada a* la función de enunciado x (o es de *longitud suficiente respecto a* x) si, y sólo si,

para todo número natural n,

si v_n aparece en x, entonces el número de lugares de f es al menos igual a n (i.e. $Np(f) \geq n$).

Definición 22b [9]

La sucesión f *satisface* la función de enunciado x si, y sólo si, f es una sucesión finita de cosas, x es una función de enunciado y

(1) f es adecuada a x

(2) x se acomoda a alguna de las cuatro condiciones siguientes:

---

[6] Las "cosas" [como aquí las denomino; podría haberlas llamado "individuos", como hace Tarski, si no fuese por el hecho de que quisiera evitar tener que mencionar la complicación, tal vez un tanto confundente, según la cual los "individuos" de Tarski resultan ser las *clases* individuales del cálculo de clases] que Tarski toma en consideración en esta sección de su obra son *clases;* a la vista del desarrollo de los §§ 4 y 5 de Tarski, hablaré aquí de "sucesiones de cosas" en lugar de hablar de sucesiones de clases, suponiendo que para todo f_t y f_k se defina una relación $f_t \subset f_k$.

[7] Cf. la Definición 6 de Tarski de la página 292 [176].

[8] *Op. cit.*, pág. 294 [178]. Tarski sólo define explícitamente la expresión "La variable v_k aparece *libre* en la función de enunciado x" [o "v_k es una variable *libre* de la función de oración x"].

[9] Es exactamente la Definición 22 [pág. 193] de Tarski excepto en que se añade

(α) Existen números naturales i y k tales que $x = \iota_{i,k}$ y $f_i \subset f_k$.

(β) Existe una función de enunciado y tal que $x = \bar{y}$, y f no satisface y.

(γ) Existen dos funciones de enunciados y y z tales que $x = y + z$ y f satisface y o z o ambas.

(δ) Existe un número natural k y una función de enunciado y tales que

(*a*) $x = P_k y$,

(*b*) toda sucesión finita g cuya longitud sea igual a f satisface y, suponiendo que g se acomode a la siguiente condición: para todo número natural n, si n es el número de un lugar de f y $n \neq k$, entonces $g_n = f_n$.

La Definición 23 de Tarski [pág. 193] se puede sustituir ahora por las dos definiciones equivalentes [10].

Definición 23+

x es un *enunciado verdadero* (i.e. x e Wr) si y sólo si (*a*) x es un enunciado ($x \in As$) y (*b*) toda sucesión finita de cosas adecuada a x satisface x.

Definición 23 ++

x es un *enunciado verdadero* ($x \in Wr$) si, y sólo si (*a*) x es un enunciado ($x \in As$) y (*b*) existe al menos una sucesión finita de cosas que satisface x.

Se puede ver que la formulación de 23 ++ no necesita hacer referencia a la adecuación de la sucesión. También se puede ver que en 23+ (que corresponde exactamente a la definición de Tarski), aunque no en 23 ++, la condición (*a*) puede sustituirse por "x es una función de enunciado", logrando así una cierta generalización al incluir funciones de enunciado con variables libres, por ejemplo, la función $\iota_{i,i}$; es decir, las funciones de enunciado universalmente válidas (*allgemeingültige* ["correctas en todo dominio individual"]) [11].

De un modo análogo, 23++ conduce a la noción de función de enunciado satisfactible (*erfüllbare*), si se extiende a las funciones.

Terminaré diciendo que en su aplicación a una teoría empírica (forma-

---

(1) a la condición de Tarski (para reemplazar sus sucesiones infinitas por otras finitas) y en que nuestro (δ) (*b*) contiene ajustes menores, en la medida en que se refiere a la longitud de f (y g). [Es desventajoso traducir "*erfüllen*" por "satisfacer", pues en la definición de "*f satisface x*" se recurre a la idea intuitiva de "x se acomoda a (i.e., satisface) tales y cuales *condiciones*". Sin embargo, ambos "satisface" son técnicamente muy distintos aunque casi coincidan intuitivamente. En la página 311 del texto alemán no se hace ninguna distinción terminológica, aunque en la nota de la página 312, correspondiente a la nota 1 de la página 193 de la traducción inglesa, se establece una distinción entre "*erfüll*" y "*befriedigt*". Evidentemente no hay ninguna circularidad en la Definición 22.]

[10] La equivalencia surge de las consideraciones de Tarski; cf. *op. cit.*, pág. 313, líneas 13 a 16 [pág. 194, líneas 12 a 15].

[11] Cf. *op. cit.*, pág. 320 [201]. Definición 27 y *sig.*

lizada al menos en parte), especialmente a funciones de enunciado no cuantificadas de dicha teoría, la definición de *cumplimiento* [o satisfacción], i.e. la Definición 22b, parece perfectamente "natural" desde un punto de vista intuitivo, sobre todo gracias a que hemos evitado las sucesiones infinitas [12].

[12] Podemos emplearla, por ejemplo, para definir la instanciación de una ley (que no está escrita como universalización; i.e., que está escrita sin prefijo universal) como una sucesión finita de cosas que la satisfacen; o, lo que es más importante en mi opinión, para definir una *instancia refutadora* de una función de enunciado (abierta o cerrada) como una sucesión finita [y adecuada] de cosas que *no* la satisfacen.

APÉNDICE

EL CUBO Y EL REFLECTOR: DOS TEORIAS ACERCA DEL CONOCIMIENTO*

El propósito de este artículo es criticar una opinión muy extendida acerca del objeto y método de las ciencias naturales, para proponer un punto de vista alternativo.

I

Empezaré haciendo una breve exposición del punto de vista que voy a examinar y que denominaré "*la teoría de la ciencia como un cubo*" (o "*la teoría de la mente como un cubo*"). El punto de partida de esta teoría viene dado por la doctrina convincente según la cual es necesario haber tenido percepciones —experiencias de los sentidos— antes de poder conocer algo acerca del mundo. Se supone que de aquí se sigue que nuestro conocimento, nuestra experiencia, consta sea de percepciones acumuladas (empirismo ingenuo), sea de percepciones asimiladas, ordenadas y clasificadas (opinión sostenida por Bacon y, en versión más radical, por Kant).

Los atomistas griegos poseían una idea un tanto primitiva de este proceso. Suponían que los átomos abandonaban los objetos percibidos para penetrar en nuestros órganos sensoriales, donde se convertían en percepciones. A partir de ellas y con el transcurso del tiempo, nuestro conocimiento del mundo externo encajaba con él perfectamente [como ocurre con las piezas de un *puzzle* que ensamblan unas en otras]. Según este punto de vista, nuestra

---

* Conferencia pronunciada (en alemán) en el Foro Europeo del Colegio Austríaco, Alpbach, Tirol, en agosto de 1948 y publicada en alemán por primera vez con el título "*Naturgesetze und theoretische Systeme*" en *Gesetz und Wirklichkeit*, editado por Simon Moser, 1949. No ha sido publicada anteriormente en inglés. [Los añadidos al texto hechos en esta traducción están entre corchetes o se indican en las notas].

El artículo anticipa muchas de las ideas desarrolladas más extensamente en este volumen y en *Conjectures and Refutations*. [Trad. castellana de Néstor Míguez, *El Desarrollo del Conocimiento Científico. Conjeturas y Refutaciones*, B. Aires, Paidos, 1967], pero también contiene otras ideas que no he publicado en ninguna parte. La mayoría de las ideas y expresiones: "la teoría de la mente como un cubo" y "la teoría de la ciencia [y de la mente] como un reflector" se retrotraen a mi estancia en Nueva Zelanda y se mencionan por primera vez, en mi *Open Society* [Trad. castellana de E. Loedel, *La Sociedad abierta y sus enemigos*, B. Aires, Paidos, 1957]. En 1946 leí un escrito sobre "La Teoría de la mente como un cubo" en el Staff Club de la London School of Economics. Este Apéndice está muy relacionado con los capítulos 2 y 5 del presente volumen.

mente es como un receptáculo —una especie de cubo— en el que se acumulan las percepciones y el conocimiento. (Bacon habla de las percepciones como si fuesen "uvas maduras y en sazón" que han de ser recogidas paciente y laboriosamente y que, una vez prensadas, destilarán el vino puro del conocimiento.)

Los empiristas extrictos nos recomiendan interferir lo menos posible con este proceso de acumulación de conocimiento. El conocimiento verdadero es el conocimiento puro no contaminado por esos prejuicios que tan proclives somos a añadir y mezclar con nuestras percepciones; sólo éstas constituyen la experiencia pura y simple. El error es el resultado de estos añadidos, de nuestras perturbaciones e interferencias con el proceso de acumulación de conocimiento. A esto Kant opone su teoría: niega que las percepciones puedan ser puras y afirma que la experiencia es el resultado de un proceso de asimilación y transformación —el resultado de combinar las percepciones de los sentidos con determinados ingredientes puestos por nuestras mentes. Las percepciones son, como si dijéramos, el material bruto que fluye del exterior del cubo, donde sufre cierta elaboración (automática) —algo semejante a la digestión o quizá a la clasificación sistemática— que termina convirtiéndolo en algo no muy distinto del "vino puro de la experiencia" de Bacon; digamos que, tal vez, en vino fermentado.

No creo que ninguno de estos dos puntos de vista nos suministre un panorama adecuado de lo que considero el proceso real de adquisición de experiencia, ni del método que se emplea en la investigación o descubrimiento. Hay que admitir que el punto de vista kantiano podría interpretarse de modo que quede más próximo a mi opinión que el empirismo puro. Naturalmente, acepto que la ciencia es imposible sin experiencia (si bien la noción de "experiencia" ha de ser cuidadosamente analizada). Aunque acepto todo esto, sostengo que las percepciones no son algo así como el material bruto —como ocurre según la "teoría del cubo"— a partir del cual construimos la "experiencia" o la "ciencia".

II

En la ciencia, lo que representa el papel esencial es la *observación* más bien que la percepción. Con todo, la observación es un proceso en el que desempeñamos un papel muy *activo*. Una observación es una percepción planificada y preparada. No "tenemos" una observación [aunque podamos "tener" una experiencia sensible], sino que "hacemos" una observación. [Los navegantes incluso "elaboran" una observación.] Las observaciones van siempre precedidas por un interés particular, una pregunta o un problema— brevemente, por algo teórico [1]. Después de todo, podemos formular toda pregunta como una hipótesis o conjetura a la que añadimos: "¿Es

[1] Con la palabra "teórico" no me refiero aquí a lo opuesto a "práctico" (puesto que nuestros intereses pueden ser muy prácticos); habría que entenderla más bien en el sentido de "especulativo" [como ocurre con un interés especulativo por un problema preexistente] frente a "perceptivo", o "racional" frente a "sensible".

así? ¿Sí o no?" Podemos, pues, decir que toda observación va precedida por un problema, una hipótesis (llámenlo como quieran); en todo caso, por algo que nos interesa, por algo teórico o especulativo. Por eso las observaciones son siempre selectivas y presuponen algo así como un principio de selección.

Antes de desarrollar más detenidamente estos puntos, intentaré introducir, a modo de digresión, algunas consideraciones de índole biológica. Aunque no pretendo que sean ni una base ni un argumento en pro de la tesis fundamental que trataré de exponer más adelante, tal vez puedan resultar útiles para superar o evitar ciertas objeciones, de modo que faciliten luego su comprensión.

III

Sabemos que todas las cosas vivas, hasta las más primitivas, reaccionan a determinados estímulos. Dichas reacciones son específicas; es decir, el número de reacciones posibles es limitado para cada organismo (y para cada tipo de organismo). Podemos decir que todo organismo posee cierto conjunto innato de reacciones posibles o cierta disposición a reaccionar de tal o cual manera. Este conjunto de disposiciones puede cambiar a medida que aumenta la edad del organismo (quizá en parte bajo la influencia de las impresiones de los sentidos o las percepciones) o bien puede permanecer constante. No obstante, sea como sea, podemos suponer que, en un momento dado de su vida, el organismo está dotado de un conjunto de posibilidades y disposiciones reactivas, conjunto que podemos considerar su estado interno [en ese momento].

De este estado interno del organismo dependerá el modo en que reaccione al medio externo. Por eso, estímulos físicos idénticos pueden producir diferentes reacciones en momentos distintos, y estímulos físicamente distintos pueden dar lugar a idénticas reacciones [2].

Así pues, sólo diremos que un organismo "*aprende de la experiencia*" si sus disposiciones reactivas cambian en el transcurso del tiempo y si podemos suponer que dichos cambios no dependen de cambios [evolutivos] internos del estado del organismo, sino que también están en función del estado cambiante del medio esterior. (Se trata de una condición necesaria, aunque no es suficiente, para poder decir que el organismo aprende de la experiencia.) En otras palabras, consideramos los procesos mediante los cuales aprende el organismo como una especie de cambio o modificación en su disposición reactiva y no, al modo de la teoría del cubo, como una acumulación (ordenada, clasificada o asociada) de recuerdos dejados por las percepciones pasadas.

Estas modificaciones en la disposición reactiva del organismo, que van a constituir el proceso de aprendizaje, se relacionan íntimamente con la no-

---

[2] Confróntese F. A. von Hayek, "Scientism and the Study of Society", *Economica, N. S., 9, 10* y *11* (1942, 1943 y 1944): [ahora también en su libro *The Counter-Revolution of Science*, 1952].

ción fundamental de "*expectativa*", así como con la de "*expectativa contrariada*". Podemos caracterizar las expectativas como *disposiciones reactivas o como preparativos para reaccionar* que se adaptan a [o anticipan] un estado futuro del medio. Esta caracterización parece más adecuada que la que describe las expectativas en términos de estados de conciencia, pues sólo tomamos conciencia de muchas de nuestras expectativas cuando se ven contrariadas al no cumplirse. Un ejemplo sería encontrar una huella inesperada en nuestro camino: el carácter inesperado de la huella nos hace conscientes del hecho de que esperábamos encontrarnos con una superficie lisa. Este contratiempo nos obliga a *corregir* nuestro sistema de expectativas. El proceso de aprendizaje consiste en gran medida en correcciones de este tipo; es decir, en la eliminación de determinadas expectativas [contrariadas].

IV

Volvamos ahora al problema de la observación. Una observación presupone siempre la existencia de un sistema de expectativas. Dichas expectativas pueden formularse en forma interrogativa, de modo que las observaciones puedan utilizarse para obtener una respuesta que confirme o corrija las expectativas así formuladas.

A primera vista, tal vez pueda haber parecido paradójica mi tesis según la cual las observaciones van precedidas por preguntas o hipótesis; pero ahora podemos ver que no es en absoluto paradójico suponer que las expectativas —es decir, las disposiciones reactivas— hayan de preceder a cualquier observación y, ciertamente, a toda percepción, pues es obvio que ciertas disposiciones o tendencias reactivas son innatas en todos los organismos, cosa que no ocurre con las percepciones y observaciones. Aunque las percepciones y, en mayor medida, las observaciones desempeñen un papel fundamental en el proceso de *modificación* de nuestras disposiciones o tendencias reactivas, es evidente que algunas de ellas han de estar presentes previamente si es que han de ser modificadas.

No se ha de pensar, en absoluto, que estas reflexiones biológicas exijan que yo adopte una posición conductista. No niego que haya percepciones, observaciones y otros estados de conciencia; lo único que hago es asignarles una función muy distinta de la que la teoría del cubo supone que desempeñan. Tampoco ha de pensarse que dichas reflexiones biológicas constituyan, en ningún sentido, un supuesto en el que han de basarse mis argumentos. Con todo, espero que nos ayuden a comprender mejor estos argumentos. Lo mismo puede decirse de las siguientes consideraciones que están muy relacionadas con estas reflexiones biológicas.

En cada instante de nuestro desarrollo pre-científico o científico vivimos en el centro de lo que acostumbro a llamar el "*horizonte de expectativas*". Con esta expresión, aludo a la suma total de nuestras expectativas conscientes, subconscientes o incluso, tal vez, enunciadas explícitamente en un lenguaje. Los animales y los bebés también poseen diversos y distintos

horizontes de expectativas, aunque, sin duda, a un nivel de conciencia más bajo que los de un científico, pongamos por caso, cuyo horizonte de expectativas consta en gran medida de teorías o hipótesis formuladas lingüísticamente.

Los diversos horizontes de expectativas difieren, evidentemente, no sólo por su mayor o menor grado de conciencia, sino también por su contenido. Con todo, en todos estos casos, el horizonte de expectativas desempeña la función de una trama de referencia: nuestras experiencias, acciones y observaciones sólo adquieren significado por su posición en esta trama.

Las observaciones, en general, tienen una función muy peculiar en esta trama. En ciertas circunstancias incluso pueden destruir la propia trama, si chocan con algunas expectativas. En tal caso, pueden tener el efecto de una bomba sobre nuestro horizonte de expectativas. Dicha bomba puede obligarnos a reconstruir o reedificar el conjunto de nuestros horizontes de expectativas; es decir, podemos vernos obligados a corregir las expectativas para hacer que encajen de nuevo en algo así como un todo consistente. Podemos decir que de este modo nuestro horizonte de expectativas se eleva y se reconstruye a un nivel superior, con lo que alcanzamos un nuevo estadio en la evolución de la experiencia; estadio en el que las expectativas que no han sido alcanzadas por la bomba se incorporan de algún modo en el horizonte, mientras que las partes dañadas se restauran o reconstruyen. Esto hay que hacerlo de modo que las observaciones perjudiciales ya no resulten disgregadoras, sino que se integren con el resto de nuestras expectativas. Si tenemos éxito en la reconstrucción, habremos creado lo que normalmente se conoce como *explicación* de esos acontecimientos observados [que han creado la desintegración, el problema].

Por lo que atañe a la relación temporal entre la observación, por un lado, y el horizonte de expectativas o teorías por el otro, podemos admitir perfectamente que una nueva explicación o hipótesis va temporalmente precedida, en general, por *aquellas* observaciones que han destruido el horizonte de expectativas precedente y que han sido el estímulo de nuestros intentos de dar con una nueva explicación. Con todo, no hemos de tomar esto como la afirmación de que las observaciones preceden siempre a las expectativas o hipótesis. Por el contrario, toda observación va precedida por expectativas o hipótesis; en especial, por aquellas que componen el horizonte de expectativas que confiere significado a dichas observaciones; sólo así alcanzan la condición de observaciones reales.

El problema, "¿Qué va primero, la hipótesis (*H*) o la observación (*O*)?", recuerda, evidentemente, aquella otra pregunta famosa: "¿Qué es primero, la gallina (*H*) o el huevo (*O*)?" Ambas cuestiones son solubles. La teoría del cubo afirma que [del mismo modo que una especie de huevo primitivo (*O*), un organismo unicelular, precede a la gallina (*H*)] la observación (*O*) va siempre antes que cualquier hipótesis (*H*), puesto que la teoría del cubo considera que estas últimas surgen de las observaciones por generalización, asociación o clasificación. Por el contrario, podemos decir ahora que la

hipótesis (expectativa, teoría o como queramos llamarla) precede a la observación, si bien la observación que refuta determinada hipótesis puede estimular otra nueva (que es, por tanto, temporalmente posterior).

Todo esto se aplica especialmente a la formación de hipótesis científicas, pues sólo mediante hipótesis aprendemos qué tipo de observaciones tenemos que hacer, hacia dónde debemos dirigir nuestra atención y en qué hemos de interesarnos. Es, por tanto, la hipótesis la que se convierte en nuestra guía y nos lleva a nuevos resultados observacionales.

Este es el punto de vista que denomino la "*teoría del reflector*" (frente a la "*teoría del cubo*"). [Según la teoría del reflector, las observaciones son secundarias respecto a las hipótesis.] No obstante, las observaciones desempeñan un papel muy importante como *contrastaciones* a que ha de someterse una hipótesis en el curso de nuestro examen [crítico] de la misma. Si la hipótesis no aprueba el examen, si queda falsada por nuestras observaciones, entonces hemos de buscar otra nueva. En tal caso, la nueva hipótesis vendrá tras aquellas observaciones que han llevado a la falsación o rechazo de la vieja hipótesis. Con todo, lo que ha conferido interés y relevancia a las observaciones y lo que al mismo tiempo ha dado pie a que las comprendiésemos inmediatamente, es la hipótesis previa, la vieja hipótesis [ahora rechazada].

De este modo, la ciencia se muestra claramente como la continuación directa del trabajo precientífico de reparación de nuestros horizontes de expectativas. La ciencia nunca parte de un punto cero, no puede estar libre de supuestos, pues en todo momento presupone un horizonte de expectativas —el horizonte de ayer, como si dijéramos—. La ciencia de hoy se construye sobre la ciencia de ayer [por lo que es el resultado del reflector de ayer] y, a su vez, la ciencia de ayer se basa en la ciencia del día anterior. Del mismo modo, las teorías científicas más antiguas se montan sobre mitos pre-científicos, los cuales se basan en expectativas aún más antiguas. Ontogenéticamente (es decir, respecto al desarrollo del organismo individual) regresamos de este modo hasta las expectativas del recién nacido; filogenéticamente (respecto a la evolución de la estirpe, del phylum) llegamos al estadio de expectativas de los organismos unicelulares. (No hay aquí ningún peligro de regresión infinita viciosa —siquiera sea porque todo organismo nace con *algún* horizonte de expectativas.) Como si dijeramos, de las ameba a Einstein no hay más que un paso.

Ahora bien, si las ciencias evolucionan de este modo, ¿Cuál es el paso característico que marca la transición de la pre-ciencia a la ciencia?

V

Aproximadamente entre el siglo cinco y seis antes de Cristo podemos encontrar en la antigua Grecia los primeros comienzos de la evolución de algo así como un método científico. ¿Qué fue lo que ocurrió allí? ¿Cuáles son los nuevos elementos en esta evolución? ¿De qué modo se relacionan

las nuevas ideas con los mitos tradicionales llegados del Este que, según creo, suministraron muchas de las sugerencias decisivas para las nuevas ideas?

Entre los babilonios y los griegos, así como entre los maorís de Nueva Zelanda —como, por otra parte, entre todos los pueblos que inventan mitos cosmológicos— encontramos narraciones acerca del comienzo de las cosas que intentan comprender o explicar la estructura del Universo en términos de la historia de sus orígenes. Dichas narraciones se hacen tradicionales y se conservan en escuelas especiales. La tradición consiste a menudo en la conservación de una clase separada o elegida, los sacerdotes o curanderos, que la guardan celosamente. Las narraciones sólo cambian poco a poco —sobre todo a merced a las imprecisiones cometidas al transmitirlas, a causa de incomprensiones y, a veces, merced a la adición de nuevos mitos inventados por profetas o poetas.

Ahora bien, lo que considero nuevo en la filosofía griega, la nueva adición a todo esto, no consiste tanto en la sustitución de los mitos por algo más "científico", cuanto en una *nueva actitud frente a los mitos*. Creo que el hecho de que su carácter empiece a cambiar no es más que una consecuencia de esta nueva actitud.

La nueva actitud a que me refiero es *la actitud crítica. En lugar de transmitir dogmáticamente la doctrina* [con el único fin de conservar la tradición auténtica] *encontramos una discusión crítica de la misma.* Algunos empiezan a plantear preguntas; ponen en tela de juicio la integridad de la doctrina: su *verdad.*

La duda y la crítica existían ya sin duda antes de este estadio. Lo nuevo, sin embargo, reside en que esa duda y crítica se convierten a su vez en parte integrante de la tradición de la escuela. Una tradición de orden superior sustituye la tradicional conservación del dogma: en lugar de la teoría tradicional —en lugar del mito— nos encontramos con la tradición de criticar teorías (que al principio difícilmente pueden ser algo más que mitos). Sólo en el transcurso de esta discusión crítica se recaba el testimonio de la observación.

No puede ser un mero accidente que Anaximandro, el discípulo de Tales, desarrollase explícita y conscientemente una teoría que se apartaba de la de su maestro ni que Anaxímenes, el discípulo de Anaximandro, se apartase de un modo igualmente consciente de la doctrina de su maestro. La única explicación plausible es que el propio fundador de la escuela desafiaba a sus discípulos a que criticasen su teoría y los discípulos convirtieron esta nueva actitud de su maestro en una tradición.

Es interesante que esto sólo haya ocurrido una vez, que yo sepa. La escuela pitagórica primitiva era sin duda del viejo tipo: su tradición no encierra la actitud crítica, sino que se limitaba a preservar la doctrina del maestro. No cabe duda de que sólo la influencia de la escuela crítica jonia relajó más tarde la rigidez de la tradición de la escuela pitagórica, preparando así el camino que llevaría al método filosófico y científico de la crítica.

La actitud de la vieja filosof'a griega encuentra su mejor expresión en las famosas líneas de Jenófanes:

Pero si los bueyes, los caballos y los leones tuviesen manos con que poder pintar
Y esculpir como hacen los hombres, entonces los caballos pintarían a sus dioses
Como caballos; los bueyes, como bueyes; todos conformarían
Los cuerpos de los dioses a imagen y semejanza de los suyos propios.

Esto no es solamente un reto crítico; es un enunciado con conciencia plena y pleno dominio de una metodología crítica.

Por tanto, creo que esta tradición de la crítica constituye una novedad característica de la ciencia. Por otro lado, me parece que la tarea que la ciencia se impone a sí misma [es decir, la explicación del mundo], así como las ideas fundamentales que utiliza, son asumidas sin romper con la construcción precientífica de mitos.

VI

¿Cuál es la misión de la ciencia? Con esta pregunta doy fin al examen preliminar de los aspectos biológicos e históricos para acceder al análisis lógico de la ciencia misma.

La función de la ciencia es en parte teórica —*explicación*— y en parte práctica —*predicción y aplicación técnica*—. Intentaré mostrar que ambas funciones son, en cierta medida, dos aspectos distintos de una y la misma actividad.

Examinaré en primer lugar la idea de explicación.

A veces se oye decir que explicar es reducir lo desconocido a lo conocido, pero no se nos dice cómo se lleva a cabo tal reducción. En cualquier caso, no es esa idea de explicación la que se utiliza en la práctica efectiva de la explicación científica. Si consideramos la historia de la ciencia para ver qué tipos de explicación eran aceptados en uno u otro momento, encontramos que en la práctica se utilizan conceptos muy diversos de explicación.

Esta mañana he expuesto en el seminario de filosofía un breve esquema de esta historia (no me refiero a la historia del concepto de explicación, sino a la historia de la práctica explicativa) [3]. Desgraciadamente, las limitaciones de tiempo me impiden tratar aquí de nuevo estas cuestiones por extenso. Con todo, mencionaré un resultado general. En el transcurso del desarro-

---

[3] (Añadido en la traducción). Parte de la historia completa se encontrará (si bien un tanto condensada y haciendo menos hincapié en lo que se ha aceptado como explicación en la práctica concreta) en mi conferencia Venecia: "Philosophy and Physics: Theories of the Structure of Matter", incluida ahora en mi libro *Philosophy and Physics* (1972). Pueden encontrarse otras partes en la primera mitad de *Conjectures and Refutations*, especialmente en los capítulos 6, 3 y 4. (Se verá que este último capítulo se solapa con algunas partes de esta conferencia, desarrollándolas.)

llo histórico de la ciencia se han considerado aceptables distintos métodos y tipos de explicación, pero todos con un rasgo común: los diversos métodos explicativos consisten en una *deducción lógica,* deducción cuya conclusión es el *explicandum* —el enunciado de la cosa a explicar— y cuyas premisas constituyen el *explicans* [el enunciado de las leyes explicativas y las condiciones]. Los cambios fundamentales que han tenido lugar en el transcurso de la historia consisten en el silencioso abandono de ciertas exigencias implícitas relativas al carácter del *explicans* (que había de ser captado intuitivamente, que había de ser autoevidente, etc.); exigencias que resultan ser irreconciliables con algunas otras cuya significación crucial se hace cada vez más clara a medida que pasa el tiempo; en particular, la exigencia de contrastabilidad independiente del *explicans* [que constituye las premisas y, por ende, el meollo mismo de la explicación].

Por tanto, una explicación es siempre la deducción del *explicandum* a partir de determinadas premisas denominadas el *explicans.*

He aquí un ejemplo un tanto truculento a modo de ilustración [4].

Ha aparecido una rata muerta y queremos saber lo que le ha sucedido. Podemos enunciar el *explicandum* de la siguiente manera: "Esta rata ha muerto recientemente". Conocemos con precisión el *explicandum,* los hechos están ante nosotros en su fría realidad. Para explicarlo, hemos de elaborar algunas explicaciones conjeturales o hipotéticas (como hacen los autores de las novelas policiacas); es decir, explicaciones que introduzcan algo *desconocido* o, en todo caso, mucho menos conocido por nosotros. Dicha hipótesis puede ser, por ejemplo, que la muerte de la rata se ha debido a la ingestión de una considerable dosis de raticida. Es una buena hipótesis en la medida en que, para empezar, nos permite formular un *explicans* a partir del cual se puede deducir el *explicandum;* en segundo lugar, nos sugiere un cierto número de contrastaciones independientes —contrastaciones del *explicans* totalmente independientes de la verdad o falsedad del *explicandum.*

Ahora bien, el *explicans* —que constituye nuestra hipótesis— no consiste solamente en la oración "Esta rata ha ingerido un cebo con una buena dosis de raticida", pues de este enunciado aislado no se puede inferir válidamente el *explicandum.* Por el contrario, hemos de utilizar como *explicans* dos tipos distintos de premisas —*leyes universales y condiciones iniciales*—. En nuestro caso, la ley universal podría formularse del siguiente modo: "Si una rata ingiere al menos 0,48 gramos de veneno, morirá en cinco minutos". La condición (singular) inicial (que es un enunciado singular) podría ser: "Esta rata ha ingerido al menos 0,48 gramos de veneno hace más de cinco minutos". De la conjunción de ambas premisas podemos deducir ahora que dicha rata ha muerto recientemente [esto es, nuestro *explicandum*].

Todo esto puede parecer bastante obvio; pero tomemos una de mis tesis —a saber, la tesis según la cual lo que he denominado las "*condiciones iniciales*" [las condiciones pertenecientes al caso individual] nunca bastan ellas

[4] En la traducción he reformado el ejemplo para que no resultase tan truculento.

solas para dar una explicación, sino que también precisamos siempre una ley general. Tal tesis no es en absoluto obvia; por el contrario, muchas veces no se admite su verdad. Incluso abrigo la sospecha de que muchos de nosotros estaríamos dispuestos a aceptar la consideración "esta rata ha ingerido raticida" como explicación suficiente de su muerte, sin añadir el enunciado explícito de la ley universal relativa a los efectos del raticida. Mas, supongamos por un momento que viviésemos en un mundo en el que cualquiera (incluso una rata) que coma una buena ración de ese producto químico denominado "raticida" se sienta especialmente bien y especialmente dichoso la semana siguiente y mucho más vigoroso que antes. Si fuese válida una ley de este tipo, ¿podríamos seguir aceptando el enunciado "Esta rata ha ingerido veneno" como explicación de su muerte? Evidentemente, no.

Por tanto, hemos llegado al resultado, con tanta frecuencia pasado por alto, de que será incompleta toda explicación que utilice solamente condiciones iniciales singulares y de que se precisa además *al menos una ley universal,* aunque en algunos casos dicha ley es tan conocida que se omite como si fuese redundante.

Resumiendo este punto: hemos descubierto que una explicación es una deducción del siguiente tipo:

U (ley universal) *I* (condiciones iniciales específicas)	} Premisas (que constituyen el *Explicans*)
E (*Explicandum*)	Conclusión

VII

¿Acaso son *satisfactorias* todas las explicaciones dotadas de esta estructura? ¿Acaso constituye nuestro ejemplo, pongamos por caso, (el ejemplo en que se explica la muerte de una rata aludiendo al matarratas) una explicación satisfactoria? No lo sabemos: las contrastaciones pueden mostrar que la rata, a pesar de haber muerto, no había ingerido un raticida.

Si un amigo se muestra escéptico acerca de nuestra explicación y nos pregunta, "¿Cómo sabes que la rata ha ingerido un veneno?", es evidente que no bastará con responder, "¿Cómo puedes dudarlo, no ves que está muerta?" Realmente, todo argumento que podemos aportar en apoyo de una hipótesis ha de ser distinto e independiente del *explicandum.* Si sólo podemos aducir como testimonio en favor de la hipótesis el *explicandum* mismo, vemos que nuestra explicación es circular y, por tanto, totalmente *insatisfactoria.* Si, por otra parte, podemos responder, "Analiza el contenido de su estómago y encontrarás una dosis de veneno" y si esta predicción (que es nueva, es decir, no está implicada por el *explicandum* sólo) muestra ser verdadera, entonces consideraremos que nuestra explicación es por lo menos una hipótesis bastante buena.

Pero he de añadir algo más, pues nuestro escéptico amigo puede poner

también en tela de juicio la verdad de la ley universal. Puede decir, por ejemplo, "De acuerdo con que esta rata ha ingerido un producto químico, ¿pero, por qué había de morir a causa de ello?". Una vez más, no podemos responder: "¿Acaso no ves que ha muerto? Esto precisamente nos muestra cuan peligroso es ingerir dicho producto", ya que ésto convertiría una vez más en circular e insatisfactoria nuestra explicación. Para hacer que sea satisfactoria habremos de someter la ley universal a contrastaciones independientes del *explicandum*.

Con esto puede considerarse por concluido mi análisis del esquema formal de explicación, pero añadiré algunas otras consideraciones y análisis al esquema general que he bosquejado.

Empezaré con una observación sobre las ideas de causa y efecto. El estado de la cuestión descrito por las *condiciones iniciales* singulares puede denominarse la "*causa*" y el descrito por el *explicandum*, el "*efecto*". Sin embargo, me da la impresión de que es preferible evitar estos términos, estando como están tan cargados de asociaciones históricas. Si a pesar de todo queremos utilizarlos, hemos de tener siempre presente que sólo adquieren un significado por respecto a una teoría o ley universal. Es la teoría o la ley la que constituye el *nexo lógico* entre la causa y el efecto, por lo que el enunciado "*A* es la causa de *B*" ha de analizarse: "Hay una teoría *T* que es contrastable y ha sido contrastada independientemente de la cual, en conjunción con una descripción, *A*, de una situación específica independientemente contrastada, podemos deducir lógicamente una descripción, *B*, de otra situación específica". (Muchos filósofos, incluso Hume, han pasado por alto la existencia de un nexo *lógico* entre "causa" y "efecto" que está presupuesto en la utilización misma de estos términos) [*].

VIII

La tarea de la ciencia no se limita a buscar explicaciones teóricas puras; también tiene aspectos prácticos: aplicaciones técnicas, así como predicciones. Ambas pueden analizarse mediante el mismo esquema lógico que hemos introducido para analizar la explicación.

(1) *La derivación de predicciones.* Mientras que en la búsqueda de una explicación el *explicandum* aparece dado —o es conocido— y hay que

[*] (Añadido en la traducción). La primera vez que hice estos comentarios sobre las nociones de "causa" y "efecto" fue en la sección 12 de mi libro *Logik der Forschung* (*The Logic of Scientific Discovery* [traducción castellana de Víctor Sánchez de Zavala, *La Lógica de la Investigación Científica*, Madrid, Tecnos, 1962]). Véase también mi libro *The Poverty of Historicism*, págs. 122 y sigs. [traducción castellana de Pedro Schwartz, *La Miseria del Historicismo*, Madrid, Taurus, 1961, págs. 150 y sigs. Reeditada en Alianza Editorial. Madrid, 1973], así como *Open Society and its Enemies* [traducción castellana de Eduardo Loedel, *La Sociedad Abierta y sus Enemigos*, Buenos Aires, Paidos, 1975], especialmente la nota 9 del capítulo 25 [nota 7 del capítulo 25 en la edición castellana citada] y "What can Logic do for Philosophy". *Aristotelian Society, Suplementary Volume, 22*, 1948, págs. 145 y sigs.

dar con un *explicans* conveniente, la derivación de prediciones procede a la inversa. Aquí aparece dada la teoría, o se supone conocida (tal vez por los libros de texto), así como las condiciones iniciales específicas (se conocen, o se supone conocidas, por observación). Lo que queda por encontrar son las consecuencias lógicas: algunas conclusiones lógicas que aún no conocemos por observación. Esto son las *predicciones*. En tal caso, la predicción *P* ocupa el lugar del *explicandum E* en nuestro esquema lógico.

(2) *Aplicación técnica.* Consideremos la tarea de construir un puente que ha de satisfacer ciertos requisitos prácticos especificados en una lista. Lo dado aquí son las especificaciones, *S*, que describen una situación requerida —el puente a construir—. (*S* son las exigencias del cliente que son distintas y anteriores a las del arquitecto.) También se nos dan las teorías físicas relevantes (incluyendo algunas reglas prácticas). Lo que hay que encontrar son ciertas condiciones iniciales a realizar técnicamente de tal carácter que de ellas, junto con la teoría, se deduzcan las especificaciones. Así, en este caso, *S* ocupa el lugar de *E* en nuestro esquema lógico *.

Esto pone de manifiesto cómo, desde un punto de vista lógico, tanto la derivación de predicciones como la aplicación técnica de teorías científicas pueden considerarse como meras inversiones del esquema básico de explicación científica.

No obstante, aún no hemos agotado la utilización de nuestro esquema: también puede servir para analizar el *procedimiento de contrastación del explicans*. El procedimiento de contrastación consiste en derivar una predicción, *P*, del *explicans* para compararla con una situación real observable. Si una predicción no está de acuerdo con la situación observada, entonces se muestra la falsedad del *explicans;* queda falsado. En tal caso, aún no sabemos si lo que es falso es la *teoría* universal o si lo que pasa es que las *condiciones iniciales* describen una situación que no corresponde a la situación real, con lo que lo falso son las condiciones iniciales. [Naturalmente, puede ocurrir perfectamente que sean falsas la teoría y las condiciones iniliales.]

Aunque la falsación de las predicciones muestre que el *explicans* es falso, no se puede sostener lo contrario: es incorrecto y muy erróneo creer que podemos interpretar la "verificación" de la predicción como "verificación" del *explicans* o de una parte de éste. La razón estriba en que de un *explicans* falso se puede deducir, válidamente, con facilidad una predicción verdadera. Incluso es confundente interpretar *toda* "verificación" o pre-

---

* (Añadido en la traducción). No ha de pensarse que este análisis entrañe que el técnico o el ingeniero se ocupe sólo de *"aplicar"* teorías suministradas por los científicos puros. Por el contrario, el técnico y el ingeniero se enfrentan constantemente con *problemas a resolver*. Dichos problemas poseen diversos grados de abstracción, pero normalmente tienen, al menos en parte, un carácter teórico. Al tratar de resolverlos, el técnico o el ingeniero, como cualquier persona, emplea el método de conjeturas o ensayos seguidos de contraste, refutación o supresión de errores. Esto se explica muy bien en la página 43 del libro de J. T. Davies, *The Scientific Approach*, 1956, libro en el que se pueden encontrar muchas ilustraciones y aplicaciones excelentes de la teoría de la ciencia como reflector.

dicción como *corroboración* práctica del *explicans*: sería más correcto decir que sólo pueden considerarse corroboraciones del *explicans* y, por tanto, de la teoría aquellas "verificaciones" o predicciones "inesperadas" [sin la teoría a examen]. Esto significa que para corroborar una teoría sólo se podrán utilizar aquellas predicciones cuya comparación con las observaciones pueda ser considerada como un intento serio de contrastar el explicans —un intento serio de refutarlo—. Podemos decir que las predicciones ["arriesgadas"] de este tipo son "relevantes para la contrastación de la Teoría" [7]. Después de todo, es perfectamente obvio que un aprobado nos dará una idea de las cualidades de un estudiante a condición de que el examen aprobado sea suficientemente riguroso. También es obvio que se pueden hacer exámenes que aprueben con facilidad [8] incluso los alumnos más flojos.

Finalmente, aparte de todo esto, nuestro esquema lógico nos permite analizar la diferencia que hay entre el objeto de las explicaciones *teóricas* y las *históricas*.

Al *teórico* le interesa encontrar y contrastar leyes universales. En el proceso de contrastación recurre a otras leyes de tipos muy diversos (a algunas de ellas de un modo totalmente inconsciente), así como a diversas condiciones iniciales específicas.

Por otro lado, *al historiador* le interesa dar con descripciones de situaciones que tienen lugar en ciertas regiones espacio-temporales finitas —es decir, lo que he denominado condiciones iniciales específicas— y contrastar o confrontar su adecuación o precisión. En este tipo de contrastes emplea, además de otras condiciones iniciales específicas, leyes universales de todo tipo —normalmente, más bien obvias— que pertenecen a su horizonte de expectativas, aunque por regla general no es consciente de que las está utilizando. En ésto se parece al teórico. [Con todo, la diferencia es muy marcada: es la diferencia que hay entre sus diversos intereses o problemas, la diferencia que hay entre lo que cada uno de ellos considera problemático.]

En un esquema lógico [similar a los anteriores] el procedimiento del teórico puede representarse del siguiente modo:

$$\begin{array}{llll} U_0 & U_0 & U_0 & \ldots \\ U_1 & U_2 & U_3 & \ldots \\ I_1 & I_2 & I_3 & \ldots \\ \hline P_1 & P_2 & P_3 & \ldots \end{array}$$

[7] En cierto sentido, una predicción relevante corresponde a una contrastación concluyente o "*experimetum crucis*", pues para que una predicción P sea relevante para contrastar una teoría T, ha de ser posible enunciar una predicción P' que no contradiga la condición inicial ni el resto de nuestro horizonte de espectativas distinto de T (suposiciones, teorías, etc.) y que contradiga a P en combinación con las condiciones iniciales y el resto del horizonte de espectativas. Esto es lo que queremos decir cuando afirmamos que P $(=E)$ ha de ser "inesperado" (sin T).

[8] Los examinadores experimentados considerarán que las palabras "con facilidad" son más bien poco realistas. Como decía algunas veces con aire meditabundo el Presidente del Consejo Gubernamental de Examinadores de Viena: "Si preguntamos en un examen, '¿cuántos son 5 más 7?' y un estudiante responde 'dieciocho', lo concede-

U_0 es aquí la ley universal, la hipótesis universal que está a examen. Se mantiene constante a través de las contrastaciones y se emplea junto con otras varias leyes U_1 U_2, ... y otras condiciones iniciales diversas I_1 I_2, ... a fin de derivar diversas predicciones P_1, P_2, ... que pueden ser confrontadas con hechos observables reales.

El procedimiento del historiador se puede representar con el esquema siguiente:

$$\begin{array}{llll} U_1 & U_2 & U_3 & \ldots \\ I_1 & I_2 & I_3 & \ldots \\ I_0 & I_0 & I_0 & \ldots \\ \hline P_1 & P_2 & P_3 & \ldots \end{array}$$

Aquí I_0 es la hipótesis histórica, la descripción histórica que ha de ser examinada o contrastada. Se mantiene constante a través de las contrastaciones y se combina con diversas leyes (sumamente obvias) U_1, U_2, U_3, ... y con condiciones iniciales correspondientes I_1, I_2, I_3, ... para derivar diversas predicciones P_1, P_2, P_3, etc.

Naturalmente, ambos esquemas están muy idealizados y excesivamente simplificados.

IX

Anteriormente, he tratado de mostrar que una explicación será *satisfactoria* sólo si sus leyes universales, su teoría, se puede constrastar independientemente del *explicandum*. Mas esto significa que una teoría explicativa satisfactoria siempre debe decir *más* de lo que ya estaba contenido en los *explicanda* que nos impulsaron inicialmente a proponerla. En otras palabras, por principio, las teorías satisfactorias deben transcender los casos empíricos que las hicieron surgir, pues de lo contrario, como hemos visto, no llevarían más que a explicaciones circulares.

Tenemos aquí un principio metodológico que está en contradicción directa con las tendencias positivistas y empiristas ingenuas [o inductivistas]. Es un principio que nos exige atrevernos a proponer hipótesis audaces (que, a ser posible, abran nuevos campos de observación) y no aquellas generalizaciones prudentes a partir de observaciones "dadas" que [desde Bacon] continúan siendo los ídolos de todo empirista ingenuo.

Nuestro punto de vista, según el cual el objeto de la ciencia es proponer explicaciones o (lo que en esencia conduce a la misma situación lógica) *

---

mos el pase. Pero cuando contesta 'verde', a veces, pienso después que realmente deberíamos de haberle dado calabazas".

* (Añadido en la traducción). En los últimos años (a partir de 1950) he establecido una distinción más tajante entre la tarea explicativa o teórica de la ciencia y la práctica o "instrumental" y he subrayado la prioridad lógica de aquella respecto de ésta. He tratado especialmente de subrayar que las predicciones no sólo poseen un aspecto instrumental, sino también y sobre todo un aspecto teórico, pues desempeñan un papel decisivo en el contraste de teorías (como ya he mostrado en esta conferencia). Véase mi libro *Conjectures and Refutations*, especialmente el capítulo 3.

crear las bases teóricas para predicciones y otras aplicaciones, este punto de vista, nos ha llevado a la exigencia metodológica de que nuestras teorías han de ser contrastables. Con todo, hay *grados de contrastabilidad*. Unas teorías se pueden contrastar *mejor* que otras. Si fortalecemos nuestra exigencia metodológica y tendemos a teorías *que se puedan contrastar cada vez mejor*, desembocamos en un principio metodológico —o un enunciado sobre el objeto de la ciencia— cuya adopción [inconsciente] en el pasado explicaría racionalmente gran número de acontecimientos de la historia de la ciencia: los explicaría como pasos que llevan al cumplimiento del objetivo de la ciencia. (Simultáneamente, enuncia dicho objeto al decirnos qué se considera *progreso* en ciencia, pues frente a la mayoría de las otras actividades humanas —particularmente en arte y música— en ciencia se da realmente un progreso.)

El análisis y comparación de los grados de contrastabilidad de diversas teorías muestra que la contrastabilidad de una teoría aumenta con su *grado de universidad*, así como con su *grado de exactitud o precisión*.

La situación es bien sencilla. Con el grado de universalidad de una teoría aumenta la amplitud de sucesos sobre los que la teoría puede hacer predicciones y, en consecuencia, también aumenta el dominio de posibles falsaciones. Ahora bien, la teoría que es más fácilmente falsada es a la vez la que mejor se puede contrastar.

Si consideramos el grado de exactitud o precisión llegamos a una situación similar. Un enunciado preciso es más fácil de refutar que otro vago y por eso puede ser mejor contrastado. Esta consideración nos permite también explicar la exigencia de que los enunciados cualitativos sean sustituidos, si ello es posible, por otros cuantitativos, en virtud de nuestro principio de aumentar el grado de contrastabilidad de las teorías. (De este modo, también podemos explicar el papel desempeñado por la *medición* en la contrastación de teorías; se trata de un recurso que se hace cada vez más importante en el transcurso del progreso científico, aunque no se puede manejar [como a menudo sucede] como rasgo característico de la ciencia o de la formación de teorías en general. No hemos de olvidar que los procedimientos de medición sólo empiezan a usarse en un estadio muy avanzado del desarrollo de algunas ciencias y que, incluso hoy, no se emplean en todas ellas. Tampoco hemos de olvidar que toda medición depende de supuestos teóricos.)

X

La transición de las teorías de Kepler y Galileo a la de Newton constituye un excelente ejemplo de la historia de la ciencia que puede utilizarse para ilustrar mi análisis.

El hecho innegable e importante de que la teoría de Newton *contradice* a ambas teorías, muestra que dicha transición nada tiene que ver con la inducción y que la teoría de Newton no se puede tomar como generalización a partir de las teorías anteriores. *Por tanto, las leyes de Kepler no se pueden*

deducir de las de Newton [a pesar de que se ha dicho frecuentemente que pueden ser deducidas a partir de ellas e incluso que éstas se pueden deducir de las de Kepler]: las leyes de Kepler sólo se pueden obtener *aproximadamente* a partir de las de Newton, aceptando la suposición [falsa] de que las masas de los diversos planetas son despreciables comparadas con la masa del sol. De modo similar, la ley de Galileo para la caída libre de los cuerpos no se puede deducir de la teoría de Newton; antes bien, la contradice. Sólo aceptando la suposición [falsa] de que la distancia total de caída es despreciable comparada con el radio terrestre podemos obtener *aproximadamente* la ley de Galileo partiendo de la teoría de Newton.

Esto muestra, naturalmente, que la teoría de Newton no puede constituir una generalización obtenida por inducción [o deducción], sino que es un nueva hipótesis que puede iluminar el camino hacia la falsación de las viejas teorías: puede iluminar y señalar la vía de acceso a esos dominios en los que, según la nueva teoría, las viejas son incapaces de arrojar buenas aproximaciones. (En el caso de Kepler se trata del dominio de la teoría de las perturbaciones y, en el de Galileo, del de la teoría de las aceleraciones variables, pues, según Newton, las aceleraciones gravitacionales varían inversamente al cuadrado de la distancia.)

Si la teoría de Newton se hubiese limitado a unir las leyes de Kepler y Galileo, habría resultado ser únicamente una *explicación circular de dichas leyes* y, por tanto, habría sido una explicación insatisfactoria. Sin embargo, su poder ilustrador y convincente estaba precisamente en su capacidad de iluminar el camino de acceso a contrastaciones independientes que nos lleva a predicciones [con éxito] incompatibles con esas teorías anteriores. Significaba una vía de acceso a nuevos descubrimientos empíricos.

La teoría de Newton es un buen ejemplo de intento de explicar determinadas teorías anteriores de grado de universalidad más bajo que no sólo conduce a una especie de unificación de dichas teorías, sino que a la vez lleva a su falsación (y por tanto, a su corrección, restringiendo o determinando el dominio en que resultan válidas como buenas aproximaciones) [10]. Tal vez sea más frecuente el caso en que la vieja teoría resulta falsada primeramente y luego surge la nueva teoría como intento de explicar el éxito parcial y el fracaso de la vieja.

XI

Creo que es importante señalar otro punto relacionado con mi análisis del concepto (o mejor la práctica) de explicación. De Descartes [tal vez incluso desde Copérnico] a Maxwell, la mayoría de los físicos trataban de explicar las nuevas relaciones descubiertas mediante *modelos mecánicos;* es

[10] (Añadido en la traducción). Pierre Duhem puso de relieve la incompatibilidad de la teoría de Newton con la de Kepler, aludiendo al "*principio de gravitación universal*" de Newton que está "*muy lejos de ser derivable por generalización e inducción a partir de las leyes observacionales de Kepler*", puesto que "*las contradice formalmente. Si es correcta la teoría de Newton, las leyes de Kepler son necesariamente fal-*

decir, intentaban reducirlas a leyes de empuje o presión con las que estaban familiarizados por el manejo cotidiano de cosas físicas —cosas pertenecientes al reino de los "cuerpos físicos de tamaño medio"—. Descartes montó sobre esta idea un programa para todas las ciencias; incluso exigía que nos limitásemos a modelos que funcionasen únicamente por empujes o presiones. Tal programa sufrió su primer derrota con el éxito de la teoría de Newton; mas su derrota (que constituía un serio motivo de aflicción para Newton y sus contemporáneos) fue pronto olvidada y la atracción gravitatoria fue admitida en el programa en pie de igualdad con el empuje y presión. También Maxwell trató de desarrollar su teoría del campo electromagnético en forma de modelo mecánico del éter, pero terminó por abandonar dicho intento. Con ello el modelo mecánico perdió casi toda su importancia: sólo quedaron las ecuaciones que originalmente pretendían describir el modelo mecánico del éter. [Se interpretaron como descripciones de ciertas propiedades no mecánicas del éter.]

Con este tránsito de una teoría mecánica a otra *abstracta,* se alcanza un estadio en la evolución de la ciencia en el que, prácticamente, sólo se exige de las teorías explicativas que puedan ser contrastadas independientemente; si ello es posible, estamos dispuestos a operar con teorías que puedan representarse intuitivamente con diagramas dibujables [o con modelos mecánicos "dibujables" o "visualizables"], que nos suministren teorías "concretas". Pero si no los podemos obtener, estamos dispuestos a trabajar con teorías matemáticas "abstractas" [que a pesar de todo pueden ser perfectamente "comprensibles" en un sentido ya analizado en otro lugar] [11].

Nuestro análisis general de la idea de explicación no se ve afectado, evidentemente, por los fallos de un modelo o representación particular. Se apli-

---

sas". (Cita de la página 193 de la traducción de P. P. Wiener del libro de Duhem, *The Aim and Structure of Physical Theory,* 1954. El término "observacional" aplicado aquí a las "*leyes de Kepler*" hay que tomarlo *cum granu salis*: las leyes de Kepler eran conjeturas tan desenfrenadas como las de Newton: no pueden inducirse a partir de las observaciones de Tycho, de mismo modo que las de Newton no pueden serlo a partir de las de Kepler.) El análisis de Duhem se basa en el hecho de que nuestro sistema solar contiene *muchos* planetas pesados a los que hay que conceder una atracción mutua según la teoría de las perturbaciones de Newton. No obstante, podemos ir más lejos que Duhem: aún cuando concedamos que las leyes de Kepler se aplican a un conjunto de *sistemas de dos cuerpos,* cada uno de los cuales contiene un cuerpo central de la masa del sol y *un* planeta (de distinta masa y distancia en los diversos sistemas pertenecientes al conjunto), aún entonces, la tercera ley de Kepler falla, si las leyes de Newton son verdaderas, como he mostrado brevemente en *Conjectures and Refutations,* nota 28 del capítulo 1, pag. 62 [traducción citada, pág. 76] y, con algún detalle, en mi artículo "El Objeto de la Ciencia" (1957), reimpreso ahora en el capítulo 5 de este volumen, así como en *Theorie und Realität,* editado por Hans Albert, 1964, capítulo 1, págs. 73 y siguientes, especialmente las págs. 82 y sigs. En este artículo hablo algo más acerca de las explicaciones que *corrigen sus explicanda* (*aparentemente "conocidos" o "dados") al explicarlos aproximadamente.* Se trata de un punto de vista que he desarrollado plenamente en conferencias a partir de 1940 (primero en un ciclo de conferencias pronunciado en la sección Cristchurch de la Royal Society de Nueva Zelanda; cf. la nota de las páginas 134 y siguiente de mi *Poverty of Historicism* [cf. la traducción citada, nota 31, págs. 161 y siguiente]).

[11] (Añadido en la traducción). En el capítulo 4 de este volumen se hace un análisis más completo de la "comprensión".

ca a todo tipo de teorías abstractas del mismo modo que se aplica a los modelos mecánicos o de otro tipo. De hecho, desde nuestro punto de vista, los modelos no son más que intentos de explicar leyes nuevas en términos de viejas leyes que ya han sido contrastadas [junto con suposiciones relativas a condiciones iniciales típicas o a la presencia de una estructura típica —es decir, el modelo en un sentido más restringido—. A menudo, los modelos desempeñan un papel importante en la extensión y elaboración de teorías, pero es preciso distinguir entre un modelo nuevo montado sobre viejas suposiciones teóricas y una nueva teoría —es decir, un nuevo sistema de supuestos teóricos.

XII

Espero que ahora parezcan menos forzadas o incluso paradójicas algunas de las afirmaciones que al principio de la conferencia podían dar esa impresión.

No hay vía ni real ni de otro tipo capaz de llevarnos necesariamente de un conjunto "dado" de hechos específicos a una ley universal. Lo que llamamos "leyes" son hipótesis o conjeturas que siempre forman parte de un sistema teórico más amplio [de hecho, de todo un horizonte de expectativas] y que, por tanto, nunca pueden ser sometidas aisladamente a contraste. El progreso de la ciencia está compuesto de ensayos, supresión de errores y ulteriores ensayos guiados por la experiencia adquirida en el transcurso de ensayos y errores previos. Nunca podemos considerar que una teoría particular es absolutamente cierta: toda teoría puede tornarse problemática por muy bien corroborada que pueda parecer ahora. Ninguna teoría científica es sacrosanta o está más allá de la crítica. Muchas veces, especialmente en el siglo pasado, hemos olvidado esto, porque estabamos impresionados por las tan repetidas, y verdaderamente magníficas, corroboraciones de determinadas teorías mecánicas que terminaron por ser consideradas indubitablemente verdaderas. El tormentoso desarrollo de la física a partir del cambio de siglo nos ha dado una buena lección; ahora hemos llegado a ver que la misión del científico es someter continuamente su teoría a nuevas contrastaciones y que ninguna teoría puede ser tenida por algo acabado. La contrastación consiste en tomar la teoría a contrastar y combinarla con todos los tipos posibles de condiciones iniciales, así como con otras teorías, para confrontar luego las predicciones resultantes con la realidad. Si desembocamos en expectativas contrariadas, en refutaciones, entonces hemos de reconstruir la teoría.

En este proceso desempeña un papel muy importante el hecho de que se vean contrariadas algunas de las expectativas con las que antaño abordábamos, ávidamente, la realidad. Puede compararse a la experiencia del ciego que choca o topa con un obstáculo, haciéndose así consciente de su existencia. *Entramos efectivamente en contacto con la "realidad" mediante la falsación de nuestras suposiciones.* La única experiencia "positiva" que

sacamos de la realidad es el descubrimiento y eliminación de nuestros errores.

Evidentemente, siempre es posible salvar una teoría falsada mediante hipótesis auxiliares [como las de los epiciclos]. Pero no es éste el camino progresivo de las ciencias. La reacción adecuada frente a una falsación es buscar teorías nuevas que parezcan ofrecernos una visión mejor de los hechos. A la ciencia no le interesa decir la última palabra, si eso significa cerrar nuestra mente a experiencias falsadoras, sino que le interesa más bien aprender de nuestra experiencia; es decir, de nuestros errores.

Hay un modo de formular las teorías científicas que apunta con particular claridad a la posibilidad de su falsación: podemos formularlas en forma de prohibiciones [o *enunciados existenciales negativos*], como por ejemplo, "No existe un sistema físico cerrado, tal que la energía cambie en una parte del mismo sin que tengan lugar cambios compensadores en otra parte" (primera ley de la termodinámica). O, "No existe una máquina con una eficiencia del 100 por 100" (segunda ley). Se puede mostrar que los enunciados universales y los existenciales negativos son lógicamente equivalentes. Esto nos permite formular todas las leyes universales del modo indicado; es decir, como prohibiciones. No obstante, se trata de prohibiciones dirigidas al técnico, no al científico. A aquél le indican cómo ha de proceder si no quiere desperdiciar sus energías. Mas para el científico son desafíos a contrastar y falsar; le incitan a intentar decubrir aquellas situaciones cuya existencia prohiben o niegan.

Hemos alcanzado, pues, un punto desde el que podemos contemplar la ciencia como una aventura apasionante del espíritu humano. Es la invención continua de teorías nuevas y el examen infatigable de su capacidad de arrojar luz sobre la experiencia. Los principios del progreso científico son muy simples. Exigen que abandonemos la vieja idea de que podemos alcanzar la certeza o incluso un alto grado de "probabilidad" en el sentido del cálculo de probabilidades con las proposiciones y teorías científicas (idea que procede de la asimilación de la ciencia con la magia y del científico con el mago): la tarea del científico no es descubrir la certeza absoluta, sino descubrir teorías cada vez mejores [o inventar reflectores cada vez más potentes] capaces de someterse a contrastaciones cada vez más rigurosas [que nos guían, por tanto, y nos desvelan siempre nuevas experiencias, iluminándolas]. Pero esto quiere decir que dichas teorías han de ser falsables: la ciencia progresa mediante su falsación.

INDICE DE AUTORES

Preparado por J. F. G. SHEARMUR

En general, las referencias a adjetivos procedentes de nombres propios aparecen citadas bajo el encabezamiento del nombre de la persona en cuestión. Las referencias de especial importancia están en *cursiva.*

Abbot, C. G., 247n*
Adler, A., 44n.
Albert, H., 40, 46n, 180n, 191, 323n.
Alcmeón, 181n.
Alexander, S., 200n, 208n.
Allison, S. K., 193n.
Anaximandro, 165n, 174 y n, 181n, 231n, 313.
Anaxímenes, 181n, 313.
Aquino, Santo Tomás de, 127.
Aristóteles, 9, 17, *23,* 98, 102n, 122 y n, 144, 148n, 152n, 176n, *184-5 y n.*

Bach, J. S., 49.
Bacon, F., *176n,* 178n, 179n, *307-8,* 320.
Baldwin, E., 224n.
Baldwin, J. M., 71, *227n, 246 y n,* 248
Barlow, N., 239n.
Bayes, T., 137.
Beethoven, L. van, *172,* 208 y n, 234-5
Bentley, R., 183n.
Bergson, H., 248, 259.
Berkeley, G., 39 y n, 42, 44, 46, 67-9 72, 88, 107-8, 124-5 y ns, 144, 165 180 n, 268.
Bernard, C., 237n.
Bertalanffy, L. von, 199n.
Black, M., 284.
Bohm, D., 201.
Bohr, N., 23, 190, 193 y n, 199n, 202 yn. 249, 269, 275.
Boltzmann, L., 273.
Bolzano, B., 76, 106, 115, 123, *124 y ns,* 144, 147, 162n, 279.
Boole, G., 136-7.
Born, M., 169n, 173-4 y n, 186n, 193, 228, 271, 274 y n.
Boscovic, R. J., 291n.
Bothe, W., 193n.
Brahe, Tycho, 323n.
Breck, A. D., 260n.
Bridgman, P., 160.
Brillouin, L., 201n.
Broglie, Prince Louis de, 201, 249.
Brouwer, L., 109-10, 116, *125-36 y ns,* 144, 277-8.
Brown, R. W., 52n.
Buda, 271.
Bühler, K., 48n, 118n, 144, 153n, 218 y n.
Bunge, M., 47n, 137, 144-6, 169n, 191, 275n.
Busch, W., 60 y n.
Butler, S., 221, 236, 248, 250.

Cajori, F., 187n.
Campbell, D. T., 71.
Carnap, R., 29 y n, 54 n, 161-2n, 266, 303n.
Carus, P., 200n.
Carus, T. L., 200n.
Catullo, 178.
César, Julio, 178.
Chadwick, Sir James, 202.
Churchill, Sir Winston, 41n, *49-50 y n,* 54n, 68, 83.
Cicerón, 17 y n, *99,* 178.
Collingwood, R. G., 151, 153n, 156, 159n, 173, *176-8n,* 178-9.
Colodny, R. G., 31n.
Compton, A. H., *Cap. 6,* 193-4 y ns, 199n, 200n, 202 y n, 203-4, 210-11, *213-19 y ns,* 222-3, 226n, 228-9, 235.
Comte, A., 241n.
Copérnico, N., 93, 163-4, 165n, 166 y n, 322.
Coriolis, G. de, 163n.
Cotes, R., 183.

Darwin, C., 30, 40, 71, 72-3, 79, 97, 117, 139, 143, 153, 209, 223-4 y n, 236-7, *239 y n,* 240, *245-9 y n,* 250, 253, 256-7 y n, 258-9, 263.

* n = nota; ns = notas.

Darwin, F., 239n.
Davies, J. T., 318n.
Descartes, R., 9, 41, *43-4*, 46 y n, 72, *76, 79*, 107, 116n, *125 y n*, 127, 144, *153n*, 183-4, *214-8 y ns*, 223, 229, 231-3 y n, 235, 290, 322-3.
Dewey, J., 67.
Diels, H., 181n.
Dilthey, W., 151, 161-2n, 162n, *165 y n*, 173, *176 y n, 178n*.
Diógenes Laercio, 150n.
Dirac, P., 193, 202, 249.
Drake, S., 166n.
Dresden, A., 129n, 144.
Duane, W., 274-5.
Ducasse, C. J., *125n*, 144.
Duhem, P., 188n, 322-3n.

Eccles, Sir John, 9, 40 y n, 75n.
Eckart, C., 204n.
Edwards, P., 53n, 290.
Einstein, A., 22, 25-6, 28, 44n, *49 y n*, *58-9 y n*, 60, 62, 64, 69 y n, 74, 85, 101-2, 126, 131, 166 y n, 169 y n, 173-4 y n, 190 y n, 193, 200-2 y ns, 209n, 228-9, 237-8 y n, 241, 243, 247, 249, 273, 301, 312.
Ellis, R. L., 179n.
Elsasser, W., 199n.
Eschenbach, M. E. von, 46-7n.
Euclides, 77, 128, 132.
Exner, F., 119n.

Feigl, H., 30n, 58n, 144, 192, *266-7*, 274n.
Fermat, P., 110.
Feuerbach, L., 41n.
Feyerabend, P. K., 30n, 58n, 138n, 144, 192, 274n.
Forster, E. M., 35.
Foucault, L., 163n.
Fowler, H. W., 54n.
Frege, G., 76, 106, *108 y n*, 109, 124 y ns, 144, 147, 149 y n, 155n.
Fresnel, A., 190 y n.
Freud, S., 46n.
Frisch, K. von, 218n, 219n.

Galileo, 28, *162-8 y ns*, 171-2, 176n, 178, 180n, 186-8 y n, *189-90*, 241 y n, 321-2.
Geiger, H., 193n.
Ghiberti, L., 170n.
Gödel, K., 53, 164n, 290-1, 295.
Goldbach, C., 116, 153.
Goldschimidt, R. B., *257-9 y ns*.
Gombrich, E. H., 132n, 145, 170-1n, 234n.
Gomperz, H., 145.
Good, I. J., 47n.
Green, D. E., 209n.

Hadamard, J., 201n.
Haldane, E. S., 144, 153.
Hamlyn, D. W., 53n.
Hanson, N. R., 202n.
Harsanyi, J. C., 29 y n.
Haydn, F. J., 171n.
Hayek, F. A. von, 112n, 116n, 145-6, 153, *174n*, 309.
Heath, D. D., 179n.
Hegel, G. W. F., 106, *122-4*, 145, 147, 148 y n, 152n, 157n, 272.
Heinemann, F., 145.
Heisenberg, W., 193, 199n, *202 y n*, 204n, 211, 216, 249, 269, 273-5.
Hempel, C. G., 31n.
Henry, P., 145.
Heráclito, 46, 122.
Herodoto, 174 y n.
Hesíodo, 174 y n.
Heyting, A., 109, *126 y n*, 134n, *134-5* y n, 145, 277-8.
Hill, E. L., 274 y n.
Hobbes, T., 9, 146, 175n.
Hochhuth, R., 41n.
Hörmann, H., 52n.
Hubel, D. H., 75n.
Humboldt, W. von, 48n.
Hume, D., 9, 16-7, *18-23 y ns*, 24-5, 27, 33-4, 35n, 37, *38*, 39 y n, 44, 46, 67, 72, 79, 83, *86-96 y ns*, 97-8, *99-101*, 107-8, 124-5, 153n *205-7 y ns*, 209n, 210-1 y n, 212-3, 229, 249, 290, 317.
Husserl, E., 145, *155n*.
Huxley, A., 207n.
Huxley, Sir Julian, 246n.

Jennings, H. S., 35 y n, *71*, 173 y n, 227n, 228n.
Jenófanes, 314.
Jessop, T. E., 144.
Jordan, P. 199n.
Jung, K., 46n.

Kahn, C. H., 231n.
Kant, I., *17n, 34*, 37-8, 42, 46, *72n*, 81n, *87 y n, 88*, 90, *92-4*, 96, 105, 107, 109, 120n, *126-9 y ns, 131-3 y ns*,

138, 145, 148n, 170n, 198n, 204n, 220, 247, 291n, 296, 307-8.
Keller, H., 154n.
Kemp-Smith, N., 126n.
Kepler, J., 28, 64, 73, 85, *164-5*, 166n, 167n, 169n, 178n, 180n, *186-8 y ns*, *188 y n*, *189-90*, 192, 197, 228, 241, 249, 271, 321, 322-3n.
Kipa (un Sherpa), 44 y n.
Kleene, S. C., 132n, 135n, 146.
Körner, S., 201n.
Kotarbinski, T., 291.
Koyré, A., 296n.
Kramers, H. A., 193 y n, 202n.
Kranz, W., 181n.
Kretschmann, E., 169n.
Kronecker, L., 116, 153.
Kuhn, T. S., 172n, 201n.

Lakatos, I., *30n*, *46n*, 117n, 124n, *132-3 y n*, 138n, 140n, 145, 158 y n, 171, 257-8 y n, 268.
Lamarck, J., 97, 139, 143, *246 y n*, *247-9*, 255, 259.
Lamettrie, J. J. (Offray de la Mettrie), 209.
Landé, A., 47n, 201, 269, 274, 275.
Laplace, P. S., 131, *200n*, 206, 208n, 213, 247.
Lashley, K. L., 52n.
Le Sage, G. L., 247 y n.
Leavis, F. R., 77n.
Leibniz, G. W., 39, 147, 198n, 291n.
Lewis, H. D., 183n.
Lillie, R., 199n, 216n.
Lindauer, M., 218n, 219n.
Livingstone, D., 283.
Locke, J., 9, 42, 44, 46, 67, 72, 79n, 88, 107-8, 124-5, 151.
Lorenz, K. Z., *73n*, 71n, 97, 228n, 248.
Luce, A. A., 144.
Lucrecio, 200n, 247n.
Lutz, B., 226n.

Macbeth, N., 258 y n.
Mach, E., 47, 88, 180n, *235n*.
MacKenna, S., 145.
Martin, R. M., 295n.
Mastin, T. H., 102n.
Marx, K., 18n, 41 y n.
Maxwell, G., 30n, 58n, 138n, 144, 192, 274n.
Maxwell, J. C., *190 y n*, 249, 276, 322-3.
Mazurkiewicz, S., 300 y n.
Medawar, Sir Peter, 237n, 245 y n, 246n.
Medicus, F., 199n.
Menger, K., 291.
Mill, J. S., 151.
Miller, D., 10, 41n, 58n, 77n.
Moore, G. E., 44 y n.
Morgan, C. L., 71.
Moser, S., 28n, 180n, 191-2, 307n.
Mostowski, A., 154n.
Motte, A., 187n.
Mozart, W. A., 208 y n, 234.
Musgrave, A., 30n, 37n, 140n, 145, 158n, 224n.
Myhill, J., 133 y n, 145.

Nagel, E., 60n, 145-6, 277n.
Neecham, J., 209n.
Nelson, E., 47n.
Neugebauer, O., 165n.
Neumann, J. von, 202n.
Neurath, O., 65 y n.
Newton, I., 22, 28, 31, 46n, 49, 58-61, 62, 64, 85, 92-3, 101-2, 128, 131, 132n, 164, 165 y n, 166, 174 y n, 180n, 183 y n, 184, *186-190 y ns*, *197-201 y ns*, 203n, 204n, 206-8, 241 y n, 243, 247 y n, 249, 276, 301, 321-2 y n, 323 y n.
Nietzsche, F., 271.
Nowell-Smith, P. H., 200n.

Ockham, William of, 266, 268, 273.
Orwell, G., 207n.

Parménides, *16*, 70n, 238n, 273.
Pauli, W., 249.
Peirce, C. S., *198-200 y n*, 201n, 203n, 206, 211, 230, 269.
Pemberton, H., 198n.
Petersen, A. F., 10.
Pirie, N. W., 209n.
Planck, M., 200 y n, 274-5.
Platón, *106*, 115, *120-2 y n*, *123-4*, 129-30, 132, 135, 145, *147-8 y n*, 156 y n, 151-2, *184-5 y n*, 272-3.
Plotino, *122 y n*, 127, 145, 152n.
Podolsky, B., 201n, 202n.
Poincaré, H., 136.
Polya, G., 129.
Powell, J. U., 175n.
Ptolomeo, 165 y n, 166n.
Pitágoras, 169n, 313.
Piteas de Marsella, 23, 98.

Quine, W. V., 65n, *266-7 y n*, 273 y n, 279.

Ramsey, F. P., 81n.
Ranke, L. von, *178n, 179n.*
Raven, C. E., 223, 237.
Reichenbach, H., 377.
Reid, T., 9, 38, 44, 72, 80, 100n.
Rembrandt, 48, 178.
Richards, I. A., 153n, 159n.
Robinson, R. M., 154n.
Rootselaar, B. van, 106n, 191.
Rosen, N., 201n, 202n.
Ross, G., 144, 153n.
Ross, Sir William, 122n.
Routledge, H., 44 y n.
Russell, B., 15, *18 y n,* 19-22, *24, 37,* 44 y n, 49 y n, 69 y n, 76, 86n, 88, 91-2, 94-5, 107-8, 109, 124n, 125 y n, 136, 161, 146.
Ryle, G., *22,* 24, 209n.

Salmon, W., *85.*
Schilpp, P. A., 49n, 69n, 202n.
Schlick, M., *200n, 210-1 y n, 212-3,* 216, 229, *282-3.*
Schopenhauer, A., 198n, 271.
Schrödinger, E., *169n,* 193, 199n, 200, 201, 204n, 228, 246 y n, 249, 268, 271, 273.
Schwartz, J., 277 y n.
Selby-Bigge, L. A., 17n, 88n, 89n, 90n, 97n, 305n.
Sexto Empírico, 99, 150n.
Shakespeare, W., 53n.
Shearmur, J., 10, 327.
Sherrington, Sir Charles, 246 y n.
Simon, A. W., 193 y n.
Simpson, G. G., 227n, 246.
Skinner, B. F., 207n.
Skolem, T. A., 291.
Slater, J. C., 193 y n, 202n.
Soddy, F., 23.
Spedding, J., 179n.
Spencer, H., 223, *Cap. 7,* 236 y n, 237n, 238, 241 y n, 242, 245, 247, 250, 256.
Spinoza, B., 42, 128, 198.
Staal, J. F., 106n, 191.
Stove, D., 22n, *91n.*
Strawson, P., *24.*
Suppes, P., 60n, 145-6, 277n.
Szilard, L., 137.

Tales, 181n, 312.
Tarski, A., 8, 48, *51, 58 y ns,* 59 y n, 63, *64-5,* 118n, 145-6, 154n, 162, 277n, *279,* 284-8, *Cap. 9 y Suplemento, 288-305 y ns.*
Teodosio, 176-8.
Teón de Esmirna, 102n.
Turing, A. M., *209n.*

Urey, H. C., 22-3.

Vesley, R., 132n, 135n, 145.
Vigier, J. P., 201.
Volkmann, R., 145.

Waddington, C. H., 227n, 246 y n.
Wald, A., 291.
Warnock, G. J., 95n.
Watkins, J. W. N., 41n, *94n,* 121n, 146, 175n.
Werde, A. van der, 45n.
Weyl, H., 174n, 273, 274 y n.
Wheeler, J. A., 228-9 y n.
Whitehead, A. N., 120.
Whorf, B. L., *131 y n,* 146, 158.
Wiener, P. P., 188n, 323n.
Wiesel, T. N., 75n.
Wigner, E. P., 47 y n, 174n.
Wisdom, J. O., 239n.
Wittgenstein, L., 124n, 146, *283 y n,* 290.
Woodger, J. H., 51n, 294n, 302n.
Wright, G. H. von, 94.

Yourgrau, W., 47n, 260n.

INDICE DE MATERIAS

Preparado por J. F. G. SHEARMUR

Los encabezamientos de particular importancia están en *cursiva;* cuando están subdivididos, los subencabezamientos se ordenan según su palabra clave. Un guión indica que se repite el subencabezamiento anterior, y un segundo guión, que se repite el sub-subencabezamiento anterior.

Abstractas: entidades, véase *mundo 3;* —y universo físico, véase *Compton (problema de);* teorías, 324.

Accidentes, 37; argumento de la improbabilidad de, 101-3.

Acción, *32-3,* 36-7, 42, 80-1, 94n, 95, 100-1, 120, 142, 143, 159, 169 y n, 178, 224, 228, 252, véase también *certeza, inducción (problema pragmático), preferencia (pragmática);* y creencia o disposición, 36-7, 48, 80-1, 94n, 99, 106-8, 110, 119-20, véase también *conocimiento (subjetivo);* y realismo, 72, 94, 98, 100.

Aceleración constante, 187-8.

Aceptación, 36, 137-8, 240, véase también *preferencia;* aceptación$_1$, aceptación$_2$ (Lakatos), 138n.

Ad hoc, 27-8, 181-3, 189-90, 248, 268, 270.

Ameba y Einstein, 35-6, 74, 228-30, 241, 243, 312.

Amplificación de efectos, 199, 216-7, y n, 231.

Analiticidad, 84-6, 92-3, 127.

Análisis situacional (o lógica —), 73, 109, 159, *168-9t* y ns,* 172-3, 175-8, véase también *situación problemática, comprensión.*

Anatómico: monstruos (Goldschmidt), 256-9; estructura (ejecutiva), 250, 252-9, véase también *evolución.*

Animal, 140, 195, conocimiento, 77, 120, 131, 142, 228-9, 240, 312, véase también *ameba* y *Einstein, supresión;* lenguaje, 117, 120, 218 y n, 219-20, véase también *lenguaje (funciones del, expresiva y señalizadora).*

Animismo, 184, 248, 250, 255.

Anti-intelectualismo, 42, 47.

Antinomias (paradojas lógicas), 53n, 54, 65, 262, 285, 289, 243.

Apariencia y realidad, 45, 47n, 273.

A posteriori, 92-3, 138-9.

Aprendizaje, 34, 44, 67, 70, 107-8, *142-4,* 239-40, 354, 269, 309-10; darwinista o evocativo *v.* lamarckiano o instructivo, 97, 139, 143, 244-6, 309-10; de los errores, 42, 175, 228, 243-4.

A priori, 18n, 37, 73, *86-7, 92-3,* 127 y n, 132n, 138-9, 213; psicológicamente (o genéticamente), 18n, 34-5, 93, véase también *innato;* validez, 18n, 35, 37-9, 86, 91-2, 97, 127-8, 140, 184.

Aproximación, 28-9, *187-90 y ns,* 247, 322-3 y n, véase también *simulación;* a la verdad, véase *verdad (aproximación a).*

Argumento, 32, 47n, 105, 110, 120, 134, *208-10, 220-2,* véase también *crítica, discusión;* y lenguaje, véase *lenguaje (funciones del, argumentadora).*

Arte, 42, 48, 106, 123, 142, *170 y n,* 172, 174, *233-5 y n,* 236.

Asociación de ideas, 17, 66-7, *71,* 79n, 91, 94-5, 207n, 310, 312, véase también *cubo, teoría del.*

Astrología, 165, 168 y n.

Atomismo, 307.

Atracción o fuerza (newtoniana), 165, 186-7, 190n, 247, 323.

Audacia, 28-9, 59, 83, 320.

Axiomatización, axiomatizabilidad, 54n, 56, 133, 154n, 297-8, 300.

Azar, 200-1 y n, 210-1, 216n, 227 y n, 229, 234, véase también *indeterminismo;* en el mundo newtoniano, *199-202 y ns,* 203n, 261.

* t = término explicado.

Baldwin, efecto, 143, 227n, 246-7 y n, *248t*, 249.
Biología, 33, 47, *98*, 202, 216-7 y n, 220, 231 y n, 237n, 242, 249, véase también *evolución;* y teoría del conocimiento, 31, 40, *69-76*, 77-80, 105, 111-4, 139-41, 240-3, 260-4, 309-12; método en, 236-7, 243-9; y percepción o psicología, 34, 67-8, 75-6, 79, 89, 97-8, 140-1, 309-10; reductibilidad de la, 265; analogías con el mundo 3, 111-6, 124, 140, 152-3, 221, 232, 200-1, 273.
Burbujas, 230-1 y n.

Cambio, realidad del, 273-4.
Cambios orientados a un fin como problema evolucionista, 161-2, 169-70, véase también *dualismo genético.*
Casos (Hume), 17-22, 97; de una ley, 32-3, 42-3.
Causalidad, *86-93*, 95-6, 113-4, 142, 171, 204 y n, 245, 271, 276, véase también *explicación;* e inducción, 86, *87*, 88-92, 95-7.
Certeza, 22-4, 31-2, 36, 45, 53-4 y n, 66-8, 78-80, 83, 98-9, 104, 127, 128, 134, 165n, 175, 242, 325; tipos de, *80-2;* búsqueda de, 45, 67, 77-80, 175, 325; *v.* comprensión, 161n, 162n, 165n, 175.
Ciencia, ciencias, 37-8, *85*, 106, 112-3, 135-6, 165n, 175, 191, 220, 240, 264-5, 277, 325; objeto o misión de la, 47, 51, 60, 83, 114, 139, 173-4, *Cap. 5*, 180-91, 242, 288 y n, 300-1, *Apéndice*, 313, 317, 321 y n; —y explicación, *180*, 182-3, 290, 240-2, 313, 321 y n; —y verosimilitud, 60-1, 62-4, 71, 300-1; aplicada, véase *tecnología;* límites de, 40n; y arte, 234n; teoría del cubo de, véase *cubo;* naturaleza conjetural de, véase *conocimiento (conjectural);* modas en, 174 y n, 200-2, 248-9; historia de, véase *historia;* y humanidades, 173-6, 264-5; normal (Kuhn), 172 y n, 201-2 y n; precursores de, 311-2, véase también *mitos;* progreso o avance en, 29, 47, 62-3, 191, 262-3, 321, 324-5, véase también *conocimiento (aumento del);* relaciones entre, 265-6; teoría de la c. como reflector, véase *reflector.*
Científico: conocimiento y teoría del conocimiento, 17, 20, 78-9, *94*, *108*, 116, 124, 126; método, véase *método;* realismo, véase *realismo;* reducción, véase *reducción.*
Cientifismo, 174 y n, 175.
Círculo de Viena, 288, 292.
Clase consecuente (sistema deductivo), véase *contenido.*
Coherencia, teoría de la verdad como, *279-80*, 281-3, *285-6.*
Comprensión, 110, *114-5*, 134, *155-79 y ns*, 261, 266-7, *272;* de una obra de arte, 170n; *v.* certeza, 161-2n, 165n, 175; teoría de Collingwood de la, 157, 175-9; grados de, 161-2 y ns; teoría de Dilthey de la, 161-2 y n, 176 y n; histórica, 156, 159, *162-70*, 174, *175-8;* en las humanidades, 155 y n, 173-6; y "problemas vivos", 171-2; ejemplo matemático de, 160-2; problema de un metaproblema, 162-2, 167-9, 171; y resolución de problemas, *157-60*, 161-79; en ciencia, 173-4, 324 y n; y análisis situacional, 168-9 y ns, 172-3, 175-9; aspectos subjetivos de la, 150, 155-9; teoría subjetiva de la, 151, 155n-7, 161-2 y ns, 165, 173, 175-9; de los organismos unicelulares, 173; teorías infructuosas, 162-9; y mundo 3, 149, 155-60, 163-5, 168, 175-9.
Complementariedad (Bohr), 247.
Compton: postulado de libertad de, 215-7, 222; problema de, 209-11, *213-14*, *215-6*, *217-8*, *222-3*, 235, véase también *mentecuerpo (problema de).*
Computadoras, 45, 77, 114, 153, *209-10 y ns*, 212, 221-2, 229-30.
Condiciones iniciales, 92, 182, 315-20, 324.
Conceptos, construcción de (Kant), 127 y n; y criterios, 289-91; teorías *v.*, 121-2 y ns, 149, 273, 280-1, 292n, véase también *esencialismo.*
Conciencia, 35-6, 106, 109, 122-4, 148, *231-3 y n*, 266, 315, véase también *conocimiento (subjetivo), estado (mental).*
Conducta: monstruosa, 257-9; mutaciones de la, véase *dualismo genético.*
Conductismo, 65, 75n, 80, 104n, 113, 149, *206-7*, 214-5, 221, *266-9*, 310.
Conjeturas: el conocimiento consta de, véase *conocimiento (conjetural);* y refutaciones, 83, 157, 160, 238-9, 292, 318n, véase también *método de ensayo.*
Conocimiento, véase también *a priori, animal básico, certeza, fundamentos,*

innato, sociología, fuentes, éxito; como conjetura, *Cap. 1,* 22-4, 33, 38-9, 47, 63, 78-82, 93, 100, 104-5, 119, 137, 139, 143, 155n, 175, 238, 242-3, 224-5, 272; demostrable, 79-80; aumento del, 42, 45, 70, *75,* 85-6, 110-1, 119-20, 123, 137-43, 261-4; véase también *ciencia (progreso de);* impredictibilidad de la —, 271; aumento del c. subjetivo, 70, 143, véase también *supresión;* objetivo, 35, *70,* 76-9, 82-3, 86, 92-4, 98, 100, *108-16,* 119-20, 142-3, 158n, 161, *271-3,* 276, véase también *mundo 3;* análogo subjetivo del —, 79-80; parcial, 212; subjetivo, 35, 48, *70-1,* 75-80, 106-11, 113, 119, 136, 155-6, 264-5, véase también *creencia, disposición, expectativas, mundo 2;* enfoque objetivo del —, 11-3, *155-60,* 168, 175-9; teoría del (epistemología), 33-6, 42-4, 50, 65-83, 86, 99-101, 103-5, *Cap. 3,* 236-43, véase también *método (científico);* — y evolución biológica, 31, 40, *69-5,* 78, 81, 105, 111-4, 139-42, 240-4, 260-4, 309-12; — y conocimiento científico, 17-22, 78-9, *94, 108,* 110, 124, 136; — subjetivo, véase *cubo, teoría del;* — tradicional, *108,* 110, 120, 125n; está impregnado de teoría, 75-6, 104; árbol del, *Cap. 7,* 241-3, 256.

Conocimiento básico, 42-3, 55, 74, 82, 157 y n, 159, 161, 163, 171, véase también *expectativas (horizonte de).*

Consecuencias inesperadas, 43, 134, 52-3 y n, véase también *mundo 3 (autonomía de).*

Construcción intuicionista, 127-30 y n, 132-6.

Contenido (clase consecuente, sistema deductivo), *56-9 y ns,* 61-4, 76-7, 83, 118-9, 150-1, 182 y n, 186, 287, *297-301;* comparabilidad de los, 56-8 y n, 59, 64; empírico, *54n,* 138-9, 245; de falsedad, *54-8 y n,* 59, 62-4, *83,* 298-302; e información, 27, 29, 31, 61; medida del, *57 y n, 58-9,* 61, 64, 101, 298, *300 y n, 301;* relativo, 55-8, 298-301; y probabilidad: 29-40, 56-7, 101, 138, 299-300; verdad, *54,* 57-8n, 61-5, 83, 119, 138, *298t-301;* cero, 55-6, 299-300.

Contenido cero, 55-6, 299-300.

Continuo, teoría del (Brouwer), 109, 134.

Contrastación, contrastabilidad, 25-7, 29, 31-2, 39, 83-4, 117, 137-9, 169, 182 y n, 238, 242, 276, 318-9 y n, 320-9, 324-5, véase también *refutación;* crucial, 26-7, 319n; grados de, 28, 30, 138, 187, 236; independiente, 28, 185-6, 313-6, 320, 322; rigurosa, 26, 30, 83, 98, 138, 181, 324; enunciados, 21, 25n, 29, 31, 39.

Contrastación, interpretación de la, 174-5 y ns, 178-9 y ns.

Control, 214-7, 222-7, 229-34; férreo, 216, 230, 234; y conciencia, 232-3, véase también *Descartes (problema de);* supresor de errores, 224 y n-7 y n; modelos de conmutador general, 215n, *216-8,* 229, 231; mediante el significado o el contenido de las teorías, 221-3, véase también *Compton (problema de);* plástico, 216-8, 222-4, 227, 229-31, 233-4.

Correspondencia: con los hechos o la realidad, 51-3, 64, 241, 265, *281-7,* 289, 292-6, 318; teoría de la verdad como, véase *verdad (teoría de la correspondencia).*

Corroboración, *29-32,* 84-5, 98, 103, 123, 190, 318, 324; grado de, 30-2, 63-4, 84-5, 103; y analiticidad de las elecciones y predicciones, 84-6; y éxito futuro, *30-1,* 324; interpretación numérica de la, 30 y n; y preferencia de teorías, 30, 85, 103; y verosimilitud, 103.

Costumbre o hábito (Hume), 91-2, *95,* 97, 206-7.

Creencia, 10, 16, 19, *33-7,* 38-9, 62, 72-3, 78-9, *80-2,* 94n, 95, 98, 100-1, 107, 110, 120 y n, 138, 196, véase también *conocimiento (subjetivo), mundo 2;* y acción, véase *acción;* eliminación de organismos con, véase *supresión;* formación de, 16, 33-4, 37, 67, *71-2,* véase también *asociación;* justificación de, véase *justificación;* y conocimiento (en la teoría del cubo), 10, 78-9, 120, 124-5 y n, 136, 141; y lógica, 108, 124 y n, 135-6, 279-80; filosofía, 35, 107, 120, 136; y probabilidad, 81 y n, 120n, 136.

Criterios, 286, 289-91, 301; de preferencia, 26; de verdad, 53, 286-7, 290-1, 301; subjetivistas, 35, 66, 68, *72,* 76.

Crítica, 30, 35, *42-3,* 94, 118-20, 123, 133, 142, 197, 238, *239-40,* 242-4, 261, 277, 287, 312; argumento o discusión —, 26, 32-3, 37, 43, 63-4, 70, *82,* 84-6, 107, 110, 119-20, 132-4, 137,

142, 157, 220, *221-3*, 262-3, 276-7, 312-3, véase también *lenguaje (funciones del, argumentadora);* actitud 39, 181, 228-9, 312-3; método, 28, 73, véase también *método de ensayo y supresión de errores;* tradiciones, 264, 312; y duda, 307; posible merced a la formulación lingüística, 35-6, 40, *70, 73*, 85, 97, 117-20, 224, 242, véase también *lenguaje (funciones del, argumentadora).*

Cuántico: postulado (Planck), 274; teoría, 30, 47, 193, 200-3 y ns, 273-8; completud, de —, 202 y n; — e indeterminismo, 200-1, 211-2, 216, 234-5; subjetivismo en —, 136, 137, 201, 275, 278.

Cubo (teoría del) de la mente (teoría de sentido común acerca del conocimiento, epistemología subjetivista), 9, 24, *Cap. 2*, 42-4 y n, 45, 49-50, *65t-73*, 75-9, 85-9, 94, 99-100, 104-5, 108, 110, 120, 124-5 y n, 133, 135-6, 140-1, 238-9 y n, 262, *Apéndice*, 307 y n, 308, 310; — *v.* realismo, *Cap. 2*, 38, 49, 69, 87-9, 99-100, 104-5, véase también *realismo;* — y religión, 68, 72, 134, 137, 183; de la ciencia, *Apéndice*, 307.

Cuerpo-mente: dualismo, 147-8, 150n, 233 y n, 250; problema, 77, 107, 142, 143, 148 y n, 149, 210, *213-17 y ns*, 222-3, 229, *231-6*, 266-9.

Darwinista, Darwinismo, 30-1, 40, 70, 73-4 y n, 79, 117, 210, 223-4, 225n, 241, 245-9, 256, 263-4, véase también *evolución;* simula el lamarckismo, 142, 227 y n, 246-7, 249; procedimiento de selección derwinista (o evocativo), 98, 139, 143, 244-6.

Decidibilidad, 289-91.

Decisiones, precipitadas, 212, 217-8.

Definición, 38, 63, 80, 122, 184n, 281-3, 294-6; definición de verdad de Tarski, véase *verdad (definición de Tarski).*

Demarcación, 15n, 25 y n, 38-9, 86.

Descartes, problema de, *314-5 y n, 216-8 y n*, 223, 232-3 y n, 235, véase también *cuerpo-mente, problema de.*

Descifrar (descodificar), *44-5, 67-8*, 69, 76, 89.

Descubrimiento, 137-41, 143, 190-2, 288; y mundo 3, 77, 116-7, 134, 148-9, 153-4 y n.

Determinismo, determinista: azar como única alternativa (Hume, Schlick), 213-6, 229; filosófico o psicológico, 304-7 y n; físico, *197 y n*, 198-9 y n, 200 y *n*, 202-5 y n, 206-7 y n, 209-13, 234; "pesadilla del —", 202-4 y n, 207-10, 234; y teorías estadísticas, 276.

Dialéctica, 121, 123, 157n, 270, 273.

Diferenciación (Spencer), 241, 245, 247.

Dios, 48; "que juega a los dados (Einstein), 173-4; (o cristianismo) y teoría del conocimiento, 68, 72, 124, 127, 183; y mundo 3, 122-3, 148n.

Discusión, 29, 37, 63-4, 70, 82-5, 119, 132-4, 137, 157, 221-2, 262, 276-7, 281, véase también *argumento, lenguaje (formulación en).*

Disposiciones, 48, 70-1, 75-9, 106-8, 111, 119-20, 154, 309-10, véase también *conocimiento (subjetivo), mundo 2;* innatas, véase *innato.*

Dogmatismo, 33, 39, 42, 68, 162, 165, 244, 249, 312-3.

Drosophila (mosca de la fruta), 256.

Dualismo: cuerpo mente, véase *cuerpo-mente;* genético, véase *genético.*

Duane, principio de, 274-5.

Einstein, teoría de, 44n, 60, 62, 64, 85, 101, 126, 174, 201, 273; y la de Newton, 22, 28, 58-9 y n, 64, 101-2, 247, 301.

Elección: de una teoría para la acción, *31-3*, 84-6, *94n;* entre teorías rivales, véase *preferencia;* secuencia de (Brouwer), 109-116.

Emergencia, 117, 119, 141, 202, 210, 221, 224-6 y n, 235; y reducción, 263-73.

Empírico/a: base, 39, 175; contenido, 54n, 55, 138-9, 245; enunciados, 25n.

Empirismo, 15, 17, 39, 67, 86, 91-2, 124-5 y n, 136, 200n, 307-8, 278; principio del, 25.

Ensayo y error, véase *método de ensayo y supresión de errores.*

Entropía, 137.

Enunciados: cálculo de, Cálculo de Sistemas, 298-9; existencial, 25n-7, 324; y hechos, véase *correspondencia.*

Enunciados básicos, 20, 25n, véase también *enunciados observacionales.*

Epistémica/s: actitudes (Ducasse), 125n; lógica, 108, 124, *136*, 279-80.

Error, 175, 228; teoría subjetivista del, 66, 72, 79n, 308.

Escepticismo, 14-5, 87, *99-101*, 105; clásico, 99, 290.
Escolasticismo, 41 y n-2.
Esencialismo, esencias, 23, 98n, 121, 182-3 y n, 184-5, 280, 292n, véase también *universales;* modificado, 184 y n, 185-6.
Espacio: intuición del, 126-7 y n, 128, 132; teoría kantiana del, 110, 126-8 y n; newtoniano, 174n.
Especialización, 172 y n, 175.
Espíritu Absoluto u Objetivo (Hegel), 106, 123-4; véase también *mundo 3 y Hegel.*
Esquema: de explicación, véase *explicación;* de ensayo y supresión de errores, véase *método de ensayo.*
Esse = *construi* (Brouwer), 129, 135, véase también *intuicionismo.*
Esse = *percipi* (Berkeley), 125, 184.
Estadística, teoría: y determinismo, 276; teoría cuántica como, 275-6.
Estado: de una discusión, 107, 215n, 233, véase también *mundo 3;* mental, 106-8, 142, 148, 150-1n, 215 y n, 232-3 y n, 266-7, 271, véase también *mundo 2;* físico, 148, 215n, 232-3 y n, 267, 271, véase también *mundo 1.*
Estímulo y respuesta, 217 y n, 224-5.
Estoicismo, 122, 147, *150 y n-1*, 290.
Estratagema convencionalista (inmunización), 39, 46n-7, 324.
Estructura, biológica, 111-3; propositiva, 116, 251, *252-9.*
Estructura de destreza, 252-60.
Estructura de disposición central, 252, 254-9, véase también *estructura propositiva, de destreza;* papel director de las mutaciones de, 254-5, 259-60, véase también *dualismo genético.*
Eter material, 190-1n, 323.
Evaluación, 63, 137-9, véase también *preferencia.*
Evidencia, grados de (Heyting), 134-5.
Evolución, 31, 38, 40, *73-5*, 85-6, 98, 111-2, 117, 119, 139-43, 202n, 209-10, 216-22, *223-8 y ns*, 229, 231-2, *Cap. 7*, 237-8, 240-2, *243-59*, 263-4, véase también *dualismo genético;* teoría general del autor, 136-7, 148; creadora (Bergson), 225, 248; y emergencia, véase *emergencia;* endosomática, 140-1, 221, 224, 232, exosomática, 97, 118, 140-1, 221-2, 224, 229, 232, 234, 260-1, véase también *mundo 3 (analogías biológicas);* problemas de los cambios dirigidos a un fin para la teoría de, 247-8, 252-7, 259-60; de los instrumentos, 241-3; y teoría del conocimiento, véase *conocimiento (teoría de y biología);* leyes de, 233, 247; carácter lógico de la teoría de la, 72-3, 223-4, 245, 247-50, 270; problemas de método en la teoría de la, 252-9; de la mente o conciencia, 231-2, 233n, 311; problemas del organismo en, véase *problemas (evolutivos);* y auto-transcendencia, 141,4; árbol de la, 241-2; impredictibilidad de, 269-70.
Exactitud, precisión, 63, 129n, 190, 203n, 206-8, 248, 321-2.
Existencial, enunciado, 25n, 26-7, 324.
Exito en la búsqueda del conocimiento: la explicación de probaría demasiado, 34, 98, 191; improbable si nuestras teorías son correctas, 24, 34, 37-8, 191.
Expectativas, 16-8, 22, *32-4*, *35-6*, 66-7, 70, 73, 75, 141-2, 239, 243, 310-2, 324, véase *conocimiento (subjetivo);* horizonte de, 35, 311t-2, 319 y n, 324, véase también *conocimiento básico;* innatas, véase *innato;* inconscientes, 34-5, 310.
Experiencia, 17-8, 25, 44, 65, 68, 104, 142, 308-9, 311, 324; y descifrado, *44-5*, *67-8*, 69, 76, 87.
Experimento: crucial, 26-7, 319n; como contrastación, 244.
Explicación, 38, 92, 120-1, *180-91*, 241-3, 245-6, 268, 255-6, *313-21*, véase también *causalidad, reducción, simulación; ad hoc,* o circular, 181-2, 190, 248, 268, 316, 320, 322; y corrección del *explicandum*, véase *explicandum;* histórica, véase *histórico;* historia de, 313, 323-4; y condiciones iniciales, 92, 182, *315-20;* de lo conocido mediante lo desconocido, 180, 313-4; y modelos, 247-8, 323-4; y predicción, 184, 313, 317-9 y n, 321; representación esquemática, 316, 319-20; del éxito en la búsqueda del conocimiento, 33, 98, 204; última, 121, 183-4, 324, véase también *esencialismo;* y leyes universales, 182, 185-6, 315-6, 518-20.
Explicandum, 108-2, 313-7, 320, véase también *explicación;* corrección de, 28, *186-92*, *322-3 y n.*
Explicans, 180-4, 313-4, 317-8, véase también *explicación.*

Falibilidad, 48, 68, 118, 130, 234, véase también *conocimiento (conjetural)*.
Falsación, véase *refutación*.
Falsedad, 25-6; 151, 220, 289; contenido, *54-8 y n*, 59, 62-4, *83, 298-301;* y proximidad a la verdad, 61-2, 287, véase también *verosimilitud;* retrasmisión de la, 40, 276-7, 279; en el "mundo 3", 149.
Fenomenalismo, 46n, 49, 204n, 269.
Fenomenología, 46n, 49, *155n*.
Fiabilidad de las teorías, 32-4, 38, 97-8, véase también *inducción (problema de, tradicional), preferencia (pragmática)*.
Filosófico: determinismo, 204-7 y n; reducción, 266-9.
Física: realismo *v.* subjetivismo en, 136-7, 273-6; reducción de la química a, 265-6, 267-8.
Fisicalismo, 210-11, 266-9.
Físico: movimiento y estados mentales, véase *Descartes, problema de;* mundo o universo, véase *mundo 1;* entidades — y abstractas, véase *Compton, problema de*.
Formalismo, 129 y n, 130, 133.
Funciones recursivas, 109, 116.
Fuentes del conocimiento, 65, 76 y n, 127-8, 130-2, véase también *cubo, teoría del*.
Función de enunciado, 301-6 y ns.
Fundamentos (base) del conocimiento, 42-3, 45, 49, 67, véase *punto de partida*.

Galileo, teoría de: y teoría de Newton, 28, 180n, *186-7 y ns*, 187-90, 141 y n, 322-3; de las mareas, *162-8 y ns*, 178-9 y n.
Geisteswissenschaften (humanidades), 151, *155, 173-80*, véase también *comprensión*.
Genéticamente *a priori*, 93, véase también *innato*.
Genético: código, 76, 143-5; dualismo (pluralismo), 140, 143-5, 224-5, 232-4, *250-9*, 253-4, 255-6.
Geometría, 93, 126, 128, 164-5n, 166n.
Gestalten, 66, 196n.
Gravedad, 22, 58-9, 67, 101, 174, *247*.

Hábito, 142-3, véase también *costumbre*.
Hermenéutica, 155 y n, 173-6*n*, 178-90n, véase también *comprensión*.
Herramienta, 221, 241; anatómicas, ejecutivas, 250, 252-9.
Heurística, 34, 113, 129, véase también *resolución de problemas*.
Hipótesis suplementarias, 324, véase también *inmunización*.
Historia, 71, 159-60, *174-9 y ns*, 261-2, 270-4; de la ciencia, 63, 102-12 y ns, 174, 313, 321, véase también *Galileo, teoría de*.
Histórica: explicación, 169, 175-9 y ns, 319-20, véase también *análisis de situaciones;* evolución y —, 247-50; comprensión, 156, 159, 162-71, 175-9, véase también *comprensión*.
Holismo, 196 y n, 197.
"Hume, problema de", véase *inducción (problema de, lógico, de Hume)*.

Idealismo, 44, *46-51*, 68-9, 72 y n, 87-8, 99n, 105, 269, 291, 295.
Ideas (Platón), 106, 121-4, 147-8, 151, 185n, 273.
Impresiones, 66-7, 72, 89, 94, 100, véase también *percepción*.
Inconscientes, teorías o expectativas, 35-6, 244, 310.
Incorporar, 144, 214-8 y ns.
Indeterminación, fórmula de Heisenberg, 200n, 217, 273-7, véase también *teoría cuántica*.
Indeterminismo (físico), *199-202 y ns*, 203n, 204n, 210-13, 216-7, 234-5; no basta, 304, *210*, 214, 216, 234, véase también *control (plástico);* y libertad, 204n, 210-4, 216-7n, 234-5; en un mundo newtoniano, 199-202 y ns, 203n; peirceano, 199-202 y ns, 229-30; y teoría cuántica, 201-2, 211-2, 217, 234-5.
Inducción, *Cap. 1*, 63, 86-103, 140-2, 186, 190, 249, 313, véase también *cubo, teoría del;* y corroboración, 85; y demarcación, 25, 38-9; y corrección del *explicandum*, 186-90, 237; y regresión infinita, 87, 92; invalidez de (o principios inductivos), 24, 38, 87, 94-5, 249; (supuesta) justificación de (o de los principios inductivos), 37-9, 91, 94-5, 98; principios o reglas de, 17, 19-20, 37-9, 87, 92-3, 98; teorías probabilísticas de la, 29, 30 y n, 31, 37, 89n, 91, 101-3, 242-3; problemas de la, *Cap. 1;* — de sentido común, 37; — lógicos: — — de Hume y su solución, *16-7 y ns*, 19-22, 24-5, 32-4, 87-97, 100, 249; reformulaciones de — —, 19-25; — — y teorías científicas uni-

versales, 17, 20-2, 94; — pragmáticos, 37-9; — psicológicos; — — de Hume y solución de Hume, 16-7, 24-5, 87-91, 95-8, 100, 249; reformulación de — —, 33-6; — tradicionales, 16, 25, 37-9, 91; crítica de — —, 16, 38-9, 91, 98-9; *soluciones del autor a los diversos —, 15-26 y ns,* 26-32, *33-8,* 91-2, 94-8; por repetición, 16-7, 33-6, 95, *98,* 100; simulada, 98, 249; infiabilidad de la llamada, 98-9.
Inductivismo, 23, 25, 27, 29 y n, 32, 320.
Inferencias: no demostrativas (Hempel), 32; válidas, 95, 276.
Inferencias inductivas, 37, 90-1, 95-7, 187.
Inmunización (Albert), 39, 46-7, 324.
Innato (psicológicamente *a priori,* etc.), véase también *objetivo;* conducta, 248, 256; creencias, 36; disposiciones, 67-70, 75, 119, 250, 309-10; expectativas, 17n, 34, 238-9; instintos, 34-5; conocimiento, 65, 93, 238-9; estructura, 75, véase también *dualismo genético.*
Instintos, 34-5.
Instrumentalismo, 68-9, 72-3, 82, 183-4, 140-2, véase también *tecnología.*
Instrumentos, 240-2; individuos como (Hegel), 123-4.
Integración (Spencer), 240-2, 245, 247.
Inteligibles, 148 y n, 158.
Interacción: mente-cuerpo, véase *mente-cuerpo, problema de;* entre los mundos 1, 2 y 3, véase también *mundo 1, mundo 2.*
Interpretación, 115, *156,* 157-8, 167n, 175, 178-9 y ns, véase también *comprensión;* de la experiencia, 44, 148n, véase también *descifrar.*
Interpretatio naturae, 178-9 y ns.
Intuición, 126 y n, *127,* 128-36.
Intuicionismo, *125-30 y ns,* 277-8.
Irracionalismo, 16-20 y ns, 24, 37, 82-3, 86n, 91-2, 95-8, 100.
Justificación: como objetivo, 38, 51; de creencia o conocimiento, 17-8, 35, 38-9, 45, 70, 84, 91, 120, 124-5, 133, 243; de la inducción o principios inductivos, 37-8, 91, 94-5, 98-9; de la preferencia de teorías, 18-9, 71, 84-6.

Kant-Laplace, hipótesis de, 247.
"Kant, problema de", véase *demarcación.*
Kantiano, kantismo, 37, 92-4; intuición, 12, *126-9 y ns,* 131-2; y física newtoniana, 92-3, 132n; filosofía de las matemáticas, 126-134 y ns; teoría del espacio, 109-10, 126-8 y n; teoría del tiempo, 126-9, 131.
Kepler, teoría y leyes, 64, 85, 164, 198, 272; inconsistencia con la teoría newtoniana, 28, 180n, *186-7 y n, 188-192,* 241, 237-8 y n.

Lamarckismo, 98, 139, 248; simulado por el darwinismo, 143, 227 y n, 246-9; simulado por el dualismo genético, 255, 259.
Laplace, demonio de, 200n, 208n.
Lenguaje, 45, 47-8, 70, 77, 112, 122n, 132, 152, 191, 256, véase también *lingüístico;* animal, 117, 120, 218 y n, 219-20; y comunicación, 107, 128-9, 132, 151, 153n, 218-9, véase también *lenguaje (funciones señalizadoras);* evolución del, 75, *153n, 218-22;* su formulación posibilita la crítica, 35-6, *40,* 70, 73, 86, 97, 117-20, 228, 244; funciones del: —argumentadora, 40, 48 y n, 73, 86, 93-4, *117-8 y n, 119-20,* 132-3, 153-4, *218n-22,* 257 *117-8 y n,* 120, 132, 153, *218n-21,* 241; carácter distinto de las —, 118n-9, 221, 227; — expresiva o sintomática, 107, 117-9, 151, 153n, 218-9, 221-2; relaciones de las — superiores con las inferiores, 118, 153n, 219, 221-3, 227; — exhortativas y persuasivas, 218n; — señalizadora o desencadenadora, 107, 117-8, 153n, *217-9 y ns,* 221-2; (o pensamiento discursivo) e intuición, 126n, 127-9, *131-5;* y matemáticas, 128-9, 132-4; meta, véase *metalenguaje;* objeto; véase *lenguaje objeto;* ordinario, 52n, *65,* 109, 285; y realidad, véase *correspondencia;* está impregnado de teoría, 39, 131, 141, 158; y "mundo 1", 150-1, 270; y "mundo 2", 77, 150-1; y "mundo 3", 115-20, 124, 133, 150-2 y n, 222-3.
Leyes: "establecidas", 22-4; falsadoras, 26; todas son hipotéticas, 22, 185, 324, véase también *conocimiento (conjetural);* caso de, 32-3, 306n; de la naturaleza, 93, 99, 174 y n, 185-6; observacional, 191-2, 323n; universales, 182, 185-6, 315-6, 319-20, 323-5.
Libertad, 122-3, *Cap. 6,* 196, 202-3, 208n, 211, 215-7 y n, 234-5, véase también *control (plástico);* postulado

de libertad de Compton, 215-7, 222; y determinismo, 303-4, 207-8 y n, 209; e indeterminismo, 202, 210-14, 216, 234-5.
Lingüística, véase también *lenguaje;* análisis, 80, 110; embrollos, 42-3; reducción, 266-9.
Lugar numérico, 304-6.
Luz, teorías de la, 190-1.

Mareas, teoría galileana de las, *162-8 y ns,* 179n.
Matemáticas, 125-6, 160-2, 276-8; profundidad en, 186; problemas epistemológicos de las, 128-9n, 130-3; desarrollo de las, 132; filosofía kantiana de las, 126 y n-7, *128-32;* lenguaje y pensamiento discursivo en, 128-9, 132-5; problemas metodológicos de las, 130, 135-6, 276-7; condición ontológica de las, 77, 116-7, *128-30 y ns,* 133-5, 153-4; y "mundo 3", 77, 116-7, 130, 149, 153-4.
Materialismo, 150n, 250, 292n; monista (fisicalismo), 210-11, 266-9, véase también *determinismo (físico).*
Medida, 202, 275, *321.*
Mental, estado, véase *estado (mental);* y movimientos físicos, véase *Descartes, problema de.*
Mente, véase *conciencia, estado (mental), mundo 2;* — cuerpo, véase *cuerpo-mente;* objetiva, *Cap. 4,* véase también *mundo 3.*
Metafísica, 37-8, 46, *47n,* 97, 185, 190-2, 292.
Metalenguaje, 52-3, 132, 284-6, 292-5 y ns.
Método: de conjetura y refutación, véase *conjetura;* científico, 17-8, 25, 33, 59, 63, 73, 83, 86, 93-4, 139, 174, 234n, 243-4, *Apéndice,* 312-3; de ensayo y supresión de errores, 28, 35-8, 67, 73, 79, 97n, 100, 130, 224, 227 y n, 233-5 y n, 249, 318n; representación esquemática del —, 117-9, 139-41, 157-8 y n, 160-1, 165-6, 224-5, 262-3, 270.
Metodología, 25, 46n, 113, 290-2, 277, 279, 289, 301, 312; problemas de en matemáticas, 129-30 y n, 135-6; reglas o principios de, 29, 39, 276-7.
Mitos, 141, 260-1, 264, 312-9.
Modelos, 163, 247-8, 323-4.
Monismo, 147; genético, 250-6; materialista, 210-11, 266-7; neutral, 87.
Monstruo, 256-9.
Monstruos prometedores, 257-9.
Mosquitos, *195-7,* 229-30.
Mundos, habría que distinguir más de tres, 107.
Mundo 1 (primero, físico, etc.), 9, 77, 85, 147, 144-9, 155, 158, 260, 292n, véase también *determinismo (físico), indeterminismo (físico), materialismo, estado (físico) sistema (físicamente cerrado); y lenguaje,* 150-1, 270; interacción con el mundo 2, 140, 148, *214-7 y ns,* 223, 231-3 y n, 235; influido por el mundo 3 (mediante el mundo 2), 111, 117, 142-3, 148-9, 151, 210-1, 213-8, 222-3, 251-2, 236; mundos 2 y 3, que algunos toman como aspectos de, 210-1, 266-7.
Mundo 2 (segundo, experiencia o pensamiento *en sentido subjetivo,* etc.), 9, 77t, 79, 81, 85, 94n, 106, 108-11, 123-5, 142, 148-51, 271-2, véase también *conciencia, disposiciones, conocimiento (subjetivo), estado (mental);* y lenguaje, 77, 150-1; y matemáticas, 129, véase también *intuicionismo;* y comprensión, 150, 155-6, 175, véase también *comprensión (teoría subjetiva de);* interacción con el mundo 1, véase *mundo 1;* como intermediario entre los mundos 1 y 3, 142-3, 148-9, 152, 232-4; enfoque tercio-mundano de los problemas, 11-3, 155-9, 168, 175-8, véase también *transferencia;* efectos del mundo 3 sobre, 77, 111, 148-9, 221-3, 232-4, 269, véase también *lenguaje (funciones del, argumentadora);* mundo 3, que algunos consideran como expresión de, 113-4, 141-2, 143, 149, 151n, 271-2.
Mundo 3 (tercer, pensamiento objetivo, especialmente *productos* de la mente humana), 9, 30, 76n, 77t-8, 81, 85, 94n, *106-25,* 133, 136, 140-4, *147-54,* 164, 221-2, 232-3, 261, 269-73, 292n, véase también *conocimiento (objetivo), lenguaje (formulación en);* autonomía de, 77, 107-8, 110-1, *114-7,* 123-4, 142-3, 151-4 y ns, 270-3; analogías biológicas de, 111-6, 124, 140, 144-5, 221, 233, 260-1, 273; Bolzano y, 106, 123-4, 147; contenido o inquilinos de, 76-7, 107, 121, 123, 134, *149-51,* 157-8, 273; descubrimiento en el, 77, 116-7, 134, 148-9, 153-4 y n; Frege y, 106, 124, 147, 149 y n; Hegel y, 123-4,

147, 148n, 152n, 273; producto humano, 115-7, 120, 130, 134, 142-3, 151-4, 272-3; y teoría del conocimiento, 106-11, véase también *conocimiento, teoría de;* y lenguaje, 115, 117, 120, 124, 133, 150 y n, 222-3; Platón y, 106, 120-4 y ns, 148-9 y ns, 151, 273; y comprensión, véase *comprensión;* influencia de sobre el mundo 1, véase *mundo 1;* — el mundo 2, véase *mundo 2;* algunos toman como expresión del mundo 2, véase *mundo 2.*
Mutación, 73, 86, 224, 225-6, 247-8, 250-9; véase *punta de flecha.*

Necesidad, causal, 87-92, 96-7, 204, 208 y n.
Neo-darwinismo, 224-5, 236-7, 248-9 y n, 250-1.
Neutralismo, neutral, monismo, 87, 291.
Newton, teoría de, 32, 50, 60-1, 92-3, 132n, 183, 249, 276, véase también *fuerza;* y determinismo, 197 y n, 198-201 y ns, 202, 206-8; y teoría de Einstein, 22, 28, 58-9 y n, 64, 101-2, 161, 301; y teoría de Galileo, 28, 164-5, 180n, *186-7 y n, 187-90,* 322-3; e indeterminismo, 198-201 y ns, 270; y teoría de Kepler, 186-7 y n, 188-90, 191-2, 322-3; refutabilidad de la, 46n.
Nicho ecológico, 117, 226, 140, 143, 226n, 259.
Nominalismo, 121-2n, véase también *universales.*
"Normal, ciencia" (Kuhn), 172n, 201-2 y n.
Nubes, *Cap, 6, 194t, 195-201 y n, 213,* 227-8, *229-31;* y relojes, véase *relojes.*

Objetivismo, 104 y n, véase también *conductismo.*
Objetivo, conocimiento, véase *conocimiento (objetivo);* modo de hablar, 22, 110, 151; objetividad, véase *lenguaje (formulación en).*
Objetivo (meta, propósito), 115, 143, 173, 180, 215, 222, 233-4, 254-5, 257, 259-60, véase también *dualismo genético;* influencia sobre las personas, véase *Compton (problema de);* de la ciencia, *Cap. 5,* véase también *ciencia (objeto de la).*
Objeto, lenguaje, 52-3, 284-6, 293-7 y ns.
Observación, 42, 69, 104, 171, 149n, 238-9, 244, 308-12, véase también *cubo, teoría del, descifrar;* carácter conjetural de, 39, 76; enunciados, 17-8 y n, 29, 31, 39; depende de una teoría o problema, 20-2, 39, 75-7, 175, 238-9, 269, 308-12.
Observacional: lenguaje, 141; leyes, 191-2, 323n.
Observador, papel del — en física, 273-6.
Ockham, navaja de, 266-7, 269, 273.
Ontología, véase *mente-cuerpo, monismo, pluralismo, mundo 1, mundo 2, mundo 3;* problemas de en — matemáticas, 128-9 y n, 130, 133-5, 153-4.
Operacionalismo, 104n-5.
Oraciones (Tarski), 51n, 52n, 288n, 291, 301.
Organo: evolución de un órgano complicado, 247-8, 164, 169, 259-60; posesión y uso de, 250-1, 253, 255-6; con autocontrol, 252, 254, 256.
Organismos: y burbujas, 230-1; evolución de multicelulares, 225, 250n, 264; y *phylum,* 225 y n, 232, 257; y controles plásticos, 224-7.
Organos ejecutivos, 250-9, véase también *evolución.*
Organos complejos, 248, 250, 255, 257-9.
Ortogénesis, 250, 257-60, véase también *dualismo genético.*

Palabras: no tienen importancia, 30; verdad o falsedad de las teorías *v.* significado de, 121-2, 280-1, véase también *esencialismo.*
Paradoja: del mentiroso, 53n, 382, 289; lógicas (antinomias), 53-6, 282-3, 285-9, 293.
Pensamiento discursivo, 126n, 127-9 y n, 131-2, 134-5.
Percepciones, 104, 120, 125 y n, 127, 148n, 239n, 307-10, véase también *cubo, teoría del,* y biología, 67-8, 75-9, 89, 97-8, 140-1, 310.
Piloto automático, 227, 251-4.
Pitagóricos, 312.
Planck, principio de, 274-5.
Platonismo, 123, 129-30, 135, 152-6 y n, véase también *esencialismo, ideas, mundo 3* y *Platón.*
Pluralismo, 147-9, 233-4, véase también *mundo 1, mundo 2, mundo 3;* y emergencia, 264-74; genético, véase *dualismo genético;* social, 196 teórico (teorías rivales), 20-1, 27-8, 58-9, 63-4, 82-5, 94-5, 119, 137-9, 142, 156-7 y n,

228n, 242-*4*, 276, 322, véase también *preferencia.*
Poder explicativo, 59, 138.
Politeísmo, 147.
Positivismo, 46n, 49, 174, 204n, 269, *290-1*, 292, 320.
Pragmático: preferencia, véase *preferencia;* problema de la inducción, 33-4, 37-8.
Pragmatista, teoría de la verdad, 279-83, 285-6.
Praxis, problemas prácticos, 158, 241, 282.
Precisión, 63, 190, 321-2; grado de, 182n; y determinismo, 203n, 206-8.
Predicción, predictibilidad, 85-6, 194, 271, 282; y explicación, 184, 313, 317-9 y n, 321; de eventos improbables, 101-3.
Preferencia (entre teorías alternativas), 19-21, *25-31*, 34, 37, 40, *58-9*, 61-3, 71, *83-6*, 94n, 95, 101-3, 107, 137-9, 157-8, 218n, 241-2, 263, 300-1; pragmática, *30-3*, 80-6, 94n, 95, véase también *acción, certeza, tecnología.*
Primer mundo, véase *mundo 1.*
Probabilidad, 40, 63-4, 324-5; y contenido, véase *contenido;* e inducción, véase *inducción (teorías probabilísticas;* relativa, 56-7, 300 y n-1; teoría subjetiva de, 81 y n, 120n, *136-7.*
Problema, problemas, 169n, 224, 272; evolutivo (de los organismos), 140-1, 224-6 y n, 234, 248, 263-4; condición tercio-mundana, 76-7, 107, 108-10, 115, 116-7, 119, 134, 148, 153-5 y n, 157-8, 160-75, 224, 272, 292n; producción *v.* estructura producida, 111-4, véase también *mundo 2;* y descubrimiento científico, 25, 139, 237-43, 262-4, véase también *método de ensayo;* desplazamiento, 158, 161, 170, 263; situaciones, 107, 109-10, 115-6, *158-8t*, 159-62, 169-70, 177-9, 200n, 225 y n, 269, 273; resolución de 137, 271, 222-3, 240-1, 243-4; — y arte, 170 y n, 232-3 y n; — y evolución, 140, 137-8, 227-8; representación esquemática de —, véase *método de ensayo y supresión de errores;* y comprensión, 157-8, *158-60*, 161-79.
Profundidad, 28, 60, *92*, *186-90 y n*, 262.
Programas de investigación, 110.
Progreso científico, véase *ciencia, progreso de.*
Proteínas, síntesis de, 144.
Proyectiles, Galileo *v.* Newton acerca de los, 186-7.
Prueba, 44-6, 81, 125, 128-36, 276-9, 290; uso de una lógica débil en las, 145-6, 278-9.
Pseudo-ciencia, 38.
Pseudo-problemas, 265.
Psicoanálisis, 46n.
Psicología, psicológico, 52n, 65-8, 86, 96 y n, 113, 149, 155 y n, 159, 174, 206-7, 214; y biología, 34, *67-8*, 75-6, 97, 309-10; determinismo, véase *determinismo (filosófico);* problema de la inducción, véase *inducción (problema de, psicológico);* y lógica, 34, 71, 79n, 97n, 108-9, 151, 271-2, véase también *transferencia, mundo 2 (efecto del mundo 3 sobre).*
Psicologismo, 86, 113, 155n, véase también *mundo 2 (mundo 3 tomado por algunos como expresión de).*
Ptolomeo, teoría de, 164-6 y ns.
Punta de lanza, teoría de las mutaciones comportamentales como, 224-5, 232, 254-5, 257, 259-60, véase también *dualismo genético.*

Química, reductibilidad a la física, 265-5, 267-8.

Racional: actividad y metas, 180; capacidad de argumentación, 47n; actitud, 39-40, 228; decisiones (*v.* precipitadas), 212, 216-7; discusión, 29, 84, 115, 264.
Racionalidad, 33-4, 36, 64, 83, 94-5, 220, 228; véase también *inducción (problema de, solución del autor);* y función argumentadora del lenguaje, 154n, 220; de las leyes de la naturaleza, 174 y n; principio de, 169; y mundo 3, 142.
Racionalismo, 266, 279.
Rata: del conductista, 206; explicación de la muerte de, 315-6; y operación de negación, 52n.
Razón, 24-5, 48, 98.
Razón suficiente, 35, 39 y n, 78-9 y n, 100; y ley del tercio excluso, 124-5 y n; para la preferencia, véase *preferencia.*
Realidad, 44-5, 64, 67, 100, 179, 184n, 186, 265, 273, 283, 292 y n, 295-6, 234, véase también *realismo;* correspondencia con, véase *correspondencia;* principio, 174n.

Realismo, *Cap. 2*, 42-4 y n, *45-51*, 64, 67-9, 87-8, 94n, 99 y n, 100, 105, 107, *Cap. 8*, 292 y n, 295-6; Churchill sobre el, 50; teoría del conocimiento del sentido común *v.* 49, 69, 87-9, 99-100, 104-5; y teoría de la correspondencia, 53, 64, 265, 282-5, *292;* Einstein sobre el, 49; y lenguaje, 47-8; y lógica, 276-8, *278-9*, 180-7; metafísica, 47 y n, 191-2, 292; y pluralismo, 264-5; ciencia y, 46-7, 99, 101, 190-1, *265*, 267, 273-8; condición del, 44n, 46-51 y n, 47 y n-51, 100-1, 104-5.

Reducción, 63, 245, 264-70, 294-5, 313, véase también *simulación;* y emergencia, 264-9, 270; lingüística o filosófica, 266-9; científica, 265-9.

Reflector, *Apéndice*, 307n, 311, 318, 322, 325.

Reflejos condicionados, 65, 71.

Refutación, refutabilidad, falsación, 25-6, 26, 32, 46-7 y ns, 76, 83, 87, 98, 117 y n, 182, *185-6*, 238, 240-2, 276, 318 y n-9, 321-5, véase también *supresión, contrastación;* inmunización frente, 39, 46n-7, 324; y realidad, 186, 295-7, 324.

Refutadora (falsadora), ley, 26.

Regresión infinita, 122, 281, 312; inducción 87, 92.

Reguladoras, ideas (principios), 46, 64, *118*, 123, 138, *210*, 223, 234, 242, 265, 281, 286.

Regularidades: creencia en o necesidad de, 34-6; pragmáticas, 23-4, 38, 91-2, 97-8, véase también *inducción (problema de, psicológico).*

Relatividad, 64, 126, 201, 223, véase también *Einstein, teoría de;* principio de Galileo, 163n.

Relativismo, 123, 279.

Religión: historia de, 174; y teoría subjetivista del conocimiento, 68, 72, 124, 127, 183.

Relojes, *Cap. 6, 194t-195*, 197-202 y n, 212-3, 229-30; "todos son nubes", *199-201 y ns, 211-2, 229-31*, véase también *indeterminismo;* "todas las nubes son", *197-9*, 200-2, 213, véase también *determinismo (físico).*

Repetición: y formación de creencias (cubo, teoría del), 33, 37, 67; inducción por, 16-7, 37, 94, 96, 98, 100.

Retroalimentación, 111-2, 116-7, 120, 122, 154, 211-2, 223, 225, 229-30, 232.

Satisfacción, 235 y n-6, 293-4, 301-2.

Schrödinger, ecuación de, 227-8, 272.

Secuela, 227n.

Segundo mundo, véase *mundo 2.*

Selección, 97, 139, 144, 221, 249, véase también *aprendizaje, selección natural.*

Selección natural, 40, 97, 139, 143-4, 223-4, 229, 237, 240, 245-6, 249, 252, véase también *evolución.*

Semántico: metalenguaje, véase *metalenguaje;* paradojas, véase *paradojas;* términos, 65, 292-6.

Semejanza, 34, 185 y n, véase también *universales.*

Sensible: datos, 61-9, 76, 104, 140-1; experiencia, 67, 94, 104, 125, 307, véase también *cubo, teoría del;* intuición (Kant), 127-8; órganos, 89, 140-1, 307.

Sentido común, *Cap. 2*, 42-5 y n, 46-7, 49, 53n, 68-70, 79,80, 98-100, 104-5, 195, 197, 269, 292; crítico, 42-7, 65, *99-101*, 103-5; y realismo, 43-4n, 45-7, 99-100, véase también *realismo;* como punto de partida, *42-3, 65*, 73, 75, 99, 292; problema de la inducción de, 37; teoría del conocimiento de, véase *cubo, teoría del; — v.* realismo, *Cap. 2*, 46, 49, 69, *87-9*, 99-100, 104-5; realismo, *Cap. 2*, 45-6, 87-9, 93-4, 99-100, 105, 292 y n, véase también *realismo; — v.* cubo, teoría del, véase *biología y conocimiento (teoría del), biología y percepción;* noción de verdad, véase *verdad (teoría de la correspondencia).*

Sentidos, 65, 88-9, 124, 238.

Señalizadora, función del lenguaje, véase *lenguaje (funciones de, señalizadora).*

Simetría y leyes de la naturaleza, 174n.

Simplicidad, 28, 138, 182n, 186.

Simulación, 143, 230-1, 244-9, 255, 259.

Sintaxis o morfología del lenguaje, 293-6 y ns.

Sistema deductivo (clase consecuente): axiomatizable, véase *contenido;* no axiomatizable, 54n, 298-9.

Sistemas, véase *relojes, nubes;* biológicos u orgánicos, 143-4, 197, 231n; cálculo de (Tarski), véase *Tarski;* abierto, 231, 235; físico, 194, 230, 252; físicamente cerrado, 202-4 y n, 210, 215n, 234, véase también *determinismo (físico).*

Situación (Collingwood), 176-9, véase también *situación problemática.*

Sociología del conocimiento, 113, 207.

Subjetivismo, 46n, 51, 79, 86, 94, 104,

134, *136-8*, 270, *274-7*, véase también *cubo, teoría del.*
Substancia, 214n, 232.
Sucesión, finita e infinita, 303-6 y ns.
Sueños utópicos, 207 y n.
Supervivencia, 31, 72-3, 223-5, 233, 241, 246, 248-9 y n, 263-4; mutaciones y monismo genético y dualismo, 252-60.
Supresión, 85, 117, 144, 160, 221-2, 227, véase también *refutación;* y evolución, 85, 224-6, 232, 263-4; del "portador" *v.* crítica, 70, 73, 86, 93-4, 97, 120, 142, 224-5, 228-9, 240; método de, 24, 144, 258-9, véase también *método de ensayo y error.*

Tabula rasa, 65t, 70 y n, 75-6, véase también *cubo, teoría del.*
Tarski: Cálculo de Sistemas de, 54-8 y n, 296-301, véase también *contenido;* definición de verdad, véase *verdad (definición de Tarski);* teoría de la verdad, véase *verdad (Teoría de Tarski de la).*
Tecnología, ciencia aplicada, 72, 82, 95, 148-9, 152, 241-2, 313, 317-8 y n, 321n.
Teleología, 114, 245-6, 248.
Teorías: rivales, véase *pluralismo;* todas son conjeturas, véase *conocimiento (conjetural);* impregnan toda percepción, conocimiento, lenguaje, etc., 20-2, 39, 75-6, 104, 131, 140-1, 158, 238-9, 308-12.
Tercer mundo, véase *mundo 3.*
Tercer reino (Leavis), 76n.
"Tercera posición", 184-6 y ns.
Tercio excluso, 125, 129, 178.
Testimonio: documental 156-7, 175, 179 y n, 180; independiente, 181-2, 191.
Tiempo, 77, 131, 174n, 257, 273-4; teoría kantiana del, *126-8 y ns*, 131-2, 134.
Todos y partes constituyentes, 196-7.
Tradición, 120, 135, 142, 264-5, 312-3.
Traducción, 283, 294.
Transferencia, principio de, *19t*, 35-6, 71-2 y n, 82, véase también *mundo 2 (aproximación tercio-mundana a los problemas del).*
Troquelado (Lorenz), 97.
Uniformidad de la naturaleza, 36, 97, 99, véase también *inducción (problema de, tradicional).*
Universales, 121-2 y ns, 184-5 y n, 186, véase también *esencias, mundo 3.*
Universalidad, grado o nivel de, 182 y n, 183-5, 190, 321.

Validez, 71, 119, 220, 222, 276, 289-90; *a priori*, véase *a priori;* de la inducción; véase *inducción (invalidez de).*
Verdad: aproximación o proximidad a la, 54, 61-4, 71, 75, 83-5, 95, 98, 103, 123, 138, 142, 191, 228n, 242-3, 247-8, 286, 298, 301, véase también *verosimilitud;* teoría de la coherencia, *279-80*, 281-3, *285-6;* teoría de la correspondencia, 51-3, 63-4, 118n, 242-3, 279, 281-7, *Cap. 9*, 292-8, véase también *verdad (definición de Tarski);* crítica de — (discutida), 282-4, 292-3; — y realismo, 53, 64, 265, 282, 286, 292; — rehabilitada por Tarski, 64, 273, 284-6, 292, 295-6; criterios de, véase *criterios;* definición de Tarski, 64-5, 285-6, 294-5 y n, 297, Suplemento del *Cap. 9;* es una ilusión (criticado), 39n, 54n, 290; en el lenguaje ordinario, 63, 285-6; no relativa al lenguaje, 52; teoría platónica de la, 149n; teoría pragmatista de la, 279-83, 285-6; de las proposiciones *v.*, veracidad de las personas, 151; como idea reguladora, 38-40, 118, 123, 220, 242-3, 265, 281, 287; búsqueda de la, 26-9, 38, 51, 54, 60-3, 95, 119, 123, 288; teoría tarskiana de la, 51-2, 64-5, 118n, 279, 284-8, *Cap. 9*, 292-8; de las teorías *v.* significado de las palabras, 121-2, 280-1, véase también *esencialismo;* atemporalidad de, 151-2; transmisión de, 40, 276-7, 279.
Verificación, 25, 124-5, 318.
Verosimilitud, 54-5 y n, *58-65*, 83, 95, 101-3, 118, 123, 138-9, 286, 298, 300n-1, véase también *verdad (aproximación a);* como meta de la ciencia, 60-3, 75, 300-1.
Vida, 24, 43, 98, 141, 204, 249n; emergencia o evolución de la, 210, 236, 265-6.
Vitalismo, 248, 255, 259.

La serie de artículos recogidos en este volumen constituye una exposición de las opiniones epistemológicas de K. R. Popper, más sistemática de lo que a primera vista pudiera parecer, dada la estructura de la obra. El autor, uno de los filósofos de la ciencia más sólidos y audaces de nuestro siglo, así como uno de los primeros críticos cualificados de la filosofía de la ciencia del positivismo lógico, expone aquí algunas de las claves fundamentales de su pensamiento.

El punto de partida de su epistemología es el realismo del sentido común, realismo crítico por cuanto que afirma constantemente —en contra del mito del marco teórico— la necesidad de poner en tela de juicio el propio punto de partida. Sin embargo, el autor rechaza la teoría del conocimiento del sentido común que concibe el conocimiento como un proceso pasivo —teoría del cubo—, consistente en acumular los datos directamente recibidos a través de los sentidos, para situar en su lugar una concepción activa del conocer —teoría del reflector—, necesariamente ligada a expectativas previas; de modo análogo a como ocurre en el proceso de la evolución biológica.

Tanto la evolución como el desarrollo del conocimiento exigen una estructura innata genéticamente a priori —aunque no válida a priori—, que suministra el material de partida que la selección natural o la crítica racional han de modificar para producir el siguiente estadio de desarrollo. Desde esta perspectiva, se lleva a cabo la crítica al inductivismo, poniendo en tela de juicio el carácter definitivo de toda pretensión de conocer, abogando por una crítica continua y sin fronteras de todas sus instancias, que no son más que conjeturas provisionales.

ISBN 84-309-0488-3

Tecnos